This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

P410 of SEQUE NO: 2; and an alanine to threonine mutation at a site corresponding to A485 of SEQ ID NO: 2

Tol

14. The DNA polymerase of claim 6 that has reduced discrimination against a non-conventional nucleotide selected from the group consisting of: dideoxynucleotides, ribonucleotides and conjugated nucleotides.

Please enter new claims 85-88 as follows:

85. The DNA polymerase of claim 10 wherein said mutation in Region II is selected from the group consisting of: a leucine to histidine mutation at a site corresponding to L408 of SEQ ID NO: 2; a leucine to phenylalanine mutation at a site corresponding to L408 of SEQ ID NO: 2; a proline to leucine mutation at a site corresponding to P410 of SEQ ID NO: 2.

35

- 86. The DNA polymerase of claim, 10, said polymerase further comprising an alanine to threonine mutation at a site corresponding to A485 of SEQ ID NO: 2.
- 87. The DNA polymerase of claim 10 that has reduced discrimination against a non-conventional nucleotide selected from the group consisting of: dideoxynucleotides, ribonucleotides and conjugated nucleotides.

Sur D

88. An isolated recombinant IDF-3 DNA polymerase that comprises an alanine to threonine mutation at a site corresponding to A485 of SEQ ID NO: 2.

REMARKS

As a result of this amendment, claims 1-3 and 5-47 and 85-87 are pending. Claims 4 and 11 and non-elected claims 48-84 are canceled without prejudice. Claims 1-3, 5, 46 and 47 are allowed. New claims 85-88 are added. New claims 85-87 are added to re-capture material removed from claims 12-14 by amendment to parent claim 10, and therefore add no new matter. New claim 88 is supported throughout the specification and therefore adds no new matter.

Objection to the Specification

The specification is objected to because page 23, line 5 recites "the conventional deoxynucleotides dATP, dCTP, dGTP and TTP." The Office Action states that it is believed that applicants intended that "TTP" be "dTTP," and requests appropriate correction. Applicants submit that TTP and dTTP are alternative ways of referring to the same molecule. The Examiner's attention is directed to the accompanying copy of page 957 of the 2001 Sigma-Aldrich chemical catalog (Exhibit A), which shows that thymidine-5'-triphosphate (Catalog No. T 0251) is correctly referred to as both dTTP and TTP. Applicants therefore request the withdrawal of the objection to the specification.

Rejection under 35 U.S.C. §112, first paragraph

Claims 6-15 are rejected under 35 U.S.C. §112, first paragraph for lack of written description. The Office Action states that *Thermococcus JDF-3* is essential to the claimed invention and that the claimed organisms are not fully disclosed nor have they been shown to be publicly available. Applicants respectfully disagree.

Applicants submit that *Thermococcus* strain JDF-3 is not essential to the claimed invention. Applicants have provided polynucleotide and amino acid sequences for wild-type JDF-3 polymerase and specified sites and substitutions for a number of mutants satisfying the limitations of claims 6-15. This is all that is required to satisfy the written description requirement. Applicants submit that it is not necessary to have access to Thermococcus strain JDF-3 in order to practice the full scope of these claims, and therefore respectfully request the withdrawal of this rejection under §112, first paragraph.

Claims 6-45 are rejected under 35 U.S.C. §112, first paragraph for lack of written description on separate grounds from the deposit issue discussed above. The Office Action states that the specification only provides representative species encompassed by the claims wherein the mutant polymerase is *Thermococcus* JDF-3 and the mutation is selected from the specified residues of SEQ ID NO: 2. The Office Action states that the mutations described "are not representative of the genus of mutations claimed which encompasses any and all mutations of any Family B or *Thermococcus* species JDF-3 polymerase which results in a decrease in 3' to 5' exonuclease activity or a reduction in discrimination against non-conventional nucleotides." The Office Action also states that there is no disclosure of any particular structure to

function/activity relationship in the claimed genuses or structural characteristic other than a decrease in 3' to 5' exonuclease activity or a reduction in discrimination against non-conventional nucleotides for which no predictability of structure is apparent, concluding that the specification fails to make clear that applicants were in possession of the claimed invention at the time of filing. Applicants respectfully disagree.

The Written Description Guidelines state that the scope of the claims determines the level of detail of the written description which must be provided to support them. The specification as originally filed must convey clearly to those skilled in the art as of its filing date, that the applicant has invented the subject matter later claimed. In order to meet the written description requirement, there must be correspondence between the language of the claims and the specification. The correspondence does not have to be literal, but must be commensurate in scope. For example, if a claim covers a genus of biological materials, then the specification must describe a sufficient number of the species within the genus to convey with reasonable clarity to those skilled in the art that the inventor was in possession of the entire genus at the time the patent application was filed, or must provide guidance to permit those skilled in the art to understand what is covered by the claims.

Under the written description requirement of §112, first paragraph, the application must be reviewed in its entirety to understand what the applicant has described as the essential features of the invention (see, e.g., Wang Labs. v. Toshiba Corp., 993 F.2d 858, 865, 26 U.S.P.Q.2d 1767, 1774 (Fed. Cir. 1993). For genus claims, the written description requirement may be satisfied through a sufficient description of a representative number of species by (a) actual reduction to practice, (b) reduction to drawings, or (c) disclosure of relevant identifying characteristics, e.g., structure of other physical/chemical properties, by functional characteristics coupled with a known or disclosed correlation between function and structure, or by a combination of these sufficient to show possession of the claimed genus. The Written Description Guidelines define a "representative number of species" as sufficient species to be representative of the entire genus. A "representative number" depends on whether one skilled in the art would recognize that the applicant was in possession of the necessary common attributes or features of the elements possessed by the genus in view of the species disclosed.

Applicants submit that the specification as a whole describes the claimed invention in terms that convey to one skilled in the art that Applicants were in possession of the claimed genus at the time the patent application was filed. First, Applicants submit that, contrary to the characterization in the Office Action, the specification describes species beyond simply Thermococcus JDF-3 polymerase mutated at specified residues of SEQ ID NO: 2. The specification provides a listing of 55 Family B polymerases and literature references describing them (Table I), as well as accession numbers for sequence information for 17 different Family B polymerases. Applicants submit that it was known in the art that multiple sequence alignment is a means of evaluating which domains of a protein are likely to have functional significance, and that such alignment had been performed for a number of Family B DNA polymerases (see, e.g., Braithwaite and Ito, 1993, Nucl. Acids Res. 21: 787-802, and Wong et al., 1988, EMBO J. 7: 37-47, cited in the specification and incorporated therein by reference; EXHIBITS B and C). Applicants submit that alignments and knowledge of conserved regions permit one skilled in the art to identify amino acids or regions in a given Family B protein that correspond to amino acids or regions identified as critical functional determinants in *Thermococcus JDF-3 DNA* polymerase. For example, the specification teaches that Family B polymerases have six conserved structural Regions, numbered I through VI (page 7, lines 20-21), and that Region II has similar structural attributes to the nucleotide binding region of Family A polymerases, including the critical positioning of a tyrosine residue (page 7, line 21 to page 8, line 4). Thus, functional information on particular regions of Thermococcus JDF-3 Family B DNA polymerase provides a correlation between structure and anticipated function of the corresponding region in each of the Family B polymerases. The specification also provides detailed methodologies to assess the impact of a given mutation on nucleotide discrimination and 3' to 5' exo activity. Applicants describe and reduce to practice mutants of JDF-3 polymerase in the subject specification (see page 51, line 9 to page 54, line 15 and Examples 1C, 1P, 1Q and Tables V and VI), thus establishing a structure/function correlation between sites to mutate and the resulting effect of those mutations on the function of the polymerase. This structure/function correlation permits the identification of regions to mutate and types of mutations to make in other Family B polymerases in order to achieve the claimed invention using other Family B polymerases. Applicants submit that the specification thus provides a number of species that is sufficient to convey to one skilled in the art that Applicants were in possession of the full scope of the

invention at the time of filing. The methodologies for assessing function of a given Family B DNA polymerase mutant also permit the skilled artisan to determine whether a given mutant fulfills the limits of the claims.

As another example of the structure/function correlations provided by the specification, the specification also describes the relationship of Family B polymerase Region III (or Motif B) to nucleotide recognition, including description of the Region III consensus sequence structure KX₃NSXYG and its functional relationship to the KX₃(F/Y)GX₂YG motif in helix O of the Family A polymerases. The specification points out that the O helix of Family A polymerases plays a role in ddNTP discrimination in Family A polymerases (page 9, line 20 to page 10, line 9). Further, the specification describes the effects of site-directed mutagenesis of Region III of VentTM polymerase and Thermococcus barosii Family B polymerase (page 10, lines 10-23).

Applicants submit that in each instance of describing the function of various domains or residues in non-JDF-3 polymerases, the specification provides a correlation between the specific residues or regions of the non-JDF-3 Family B polymerases and the corresponding residues or regions in the JDF-3 Family B DNA polymerase. For example, in describing mutagenesis studies on VentTM polymerase, the specification states that the studies "targeted an alanine analogous to A485 of the Thermococcus species JDF-3 DNA polymerase" (page 10, lines 10-11). As another example, this time referring to mutagenesis studies of Region II of VentTM polymerase, the specification states "site directed mutagenesis of VENTTM DNA polymerase demonstrated that three mutations at Y412 (*which corresponds to JDF-3 DNA polymerase Y409*) could alter nucleotide binding" (page 9, lines 15-17; emphasis added). Thus, Applicants submit that the specification provides structure/function relationships for regions of Family B DNA polymerases that are reasonably applicable to all Family B polymerases.

Applicants submit that the structure/function relationships described in the specification between regions of other polymerases (both Family B and non-Family B) and regions of the JDF-3 Family B DNA polymerase, combined with knowledge in the art regarding alignment and correspondence of aligned regions, are sufficient to convey to one skilled in the art that Applicants were in possession of the claimed invention. Specifically, the specification describes 25 mutant clones of exo⁻ JDF-3 Family B DNA polymerase (representing at least 4 different individual mutations covering both Regions II and III) and their relative nucleotide discrimination (see Tables V and VI). These mutants, together with the provided description of

structure/function relationships of Family B Regions II and III, the provided methods of testing mutants for reduced discrimination, and knowledge in the art regarding alignments and functional regions of Family B DNA polymerases provide adequate written description for the genus of isolated recombinant Family B DNA polymerases having reduced discrimination against non-conventional nucleotides wherein the polymerase has a mutation in Region II, as claimed in independent claim 10. For the same reasons, the specification provides adequate written description of the genus of recombinant Family B DNA polymerases from Thermococcus species JDF-3 that are 3' to 5' exonuclease deficient as claimed in independent claim 6. Also for the same reasons, the specification satisfies the written description requirement for the genus of isolated recombinant Family B DNA polymerases comprising an alanine to threonine mutation at the site corresponding to A485 of SEQ ID NO: 2 or a mutation at a site corresponding to L408, S345 or P410 of SEQ ID NO: 2, wherein the DNA polymerase has reduced discrimination against non-conventional nucleotides relative to the wild-type form of that polymerase, as claimed in independent claim 16. Because these independent claims are described in a manner meeting the written description requirement with respect to the genus claimed, it follows that the requirement is also satisfied with regard to their dependent species claims.

With regard to the scope of written description, Applicants wish to finally note that the claims reciting particular amino acid positions for mutation (e.g., claims 7-9 and 12-13, among others) recite not specific mutations of the JDF-3 Family B DNA polymerase of SEQ ID NO: 2, but mutations of amino acids "corresponding to" specific amino acids thereof. These claims therefore recognize and set forth the principle that corresponding structures or amino acids from this JDF-3 Family B DNA polymerase are applicable to other Family B DNA polymerases. Therefore, Applicants submit that the specification, of which the originally filed claims are a part, clearly describes Family B DNA polymerases in a scope broader than specific mutants of SEQ ID NO: 2, and in fact broad enough to encompass the full scope of the claims.

If the Examiner does not agree that applicants have described a sufficient number of species to fulfill the written description requirement for the genus, Applicants respectfully request that the Examiner provide Applicants with a number of species that would fulfill the written description requirement for the genus and to provide citations to case law supporting the Examiner's position. In view of the above, Applicants respectfully request the withdrawal of the written description rejection of claims 6-45.

Rejection under 35 U.S.C. §112, second paragraph

Claims 10-45 are rejected under 35 U.S.C. §112, second paragraph as being indefinite for use of the term "non-conventional nucleotides." The Office Action states that "while it is clear that dATP, dCTP, dGTP and (d)TTP are considered to be 'conventional nucleotides' it is unclear what other nucleotides, if any are also considered to be "conventional." Applicants respectfully disagree.

Applicants submit that the specification defines "non-conventional nucleotide" on page 25 of the specification as referring to "a) a nucleotide structure that is not one of *the four* conventional deoxynucleotides dATP, dCTP, dGTP and TTP recognized by and incorporated by a DNA polymerase, b) a synthetic nucleotide that is not one of *the four* conventional deoxynucleotides in (a), c) a modified conventional nucleotide, or d) a ribonucleotide (since they are not normally recognized or incorporated by DNA polymerases) and modified forms of a ribonucleotide" (emphasis added). Applicants submit that this definition makes it clear that according to the invention and with respect to DNA polymerases, there are *only* four "conventional" nucleotides, namely dATP, dCTP, dGTP and TTP. In referring to "*the four* conventional nucleotides," the definition leaves no room for more than those four conventional deoxynucleotides listed. Therefore, there is no ambiguity in the term "non-conventional nucleotides" as it (and the conventional nucleotides) is defined in the specification. Applicants respectfully request that the §112, second paragraph rejection of claims 10-45 be withdrawn.

Rejection under 35 U.S.C. §102(e)

Claims 10, 11, 14, 15 and 44 are rejected under 35 U.S.C. §102(e) as anticipated by Riedl et al., U.S. Patent No. 5,882,904. The Office Action states that Riedl et al. teaches a mutant Thermococcus barossi DNA polymerase with reduced 3' to 5' exonuclease activity and reduced discrimination against dideoxynucleotides or ribonucleotides relative to the wild type. Applicants respectfully disagree.

Applicants submit that Riedl et al. does not teach an isolated recombinant Family B DNA polymerase having reduced discrimination against non-conventional nucleotides, wherein the DNA polymerase has a mutation in Region II, as required by amended claim 10. Specifically, Riedl et al. does not teach a Family B DNA polymerase mutated in Region II. Applicants submit

that the language of the amendment "wherein the DNA polymerase has a mutation in Region II" is supported in the specification at, for example, page 52, lines 19-23. The limits of the regions of the Family B DNA polymerases, including Region II are set out in the Braithwaite and Ito and Wong et al. references (Exhibits A and B) cited in the specification and incorporated in the specification by reference. The central consensus element of Region II of the Family B DNA polymerases is also described in the specification at page 52, line 20.

Applicants submit that Riedl et al. teaches only *Thermococcus barosii* Family B DNA polymerase mutants bearing mutations in Region III, covering amino acids 488-493, and the exo mutation at amino acids 141 and 143. The reference does not teach any mutation in Region II, as required by claim 10 as amended. Therefore, applicants submit that Riedl et al. does not anticipate the invention of claim 10 and its dependents 11, 14, 15 and 44. Applicants respectfully request withdrawal of the §102(e) rejection as applied to these claims.

In view of the above, Applicants submit that all issues pertinent to patentability raised in the Office Action have been addressed herein. Applicants therefore respectfully request reconsideration of the claims.

Respectfully submitted,

Date:

1/15/02

Kathleen M. Williams Registration No. 34,380 Palmer & Dodge, LLP

111 Huntington Avenue at Prudential Center

Boston, MA 02199-7613 Tel: (617) 239-0451

Fax: (617) 227-4420

Version of amended claims marked to show changes:

- 6. (Amended) An isolated recombinant <u>Family B</u> DNA polymerase from *Thermococcus* species JDF-3 that is 3' to 5' exonuclease deficient.
- 10. (Amended) An isolated recombinant <u>Family B DNA</u> polymerase having reduced discrimination against non-conventional nucleotides, <u>wherein said DNA polymerase has a mutation in Region II.</u>
- 12. (Amended) The DNA polymerase of claim 6 [or 10] wherein said DNA polymerase further comprises a mutation selected from the group consisting of: a leucine to histidine mutation at a site corresponding to L408 of SEQ ID NO: 2; a leucine to phenylalanine mutation at a site corresponding to L408 of SEQ ID NO: 2; a proline to leucine mutation at a site corresponding to P410 of SEQ ID NO: 2; and an alanine to threonine mutation at a site corresponding to A485 of SEQ ID NO: 2.
- 14. (Amended) The DNA polymerase of claim 6 [or 10] that has reduced discrimination against a non-conventional nucleotide selected from the group consisting of: dideoxynucleotides, ribonucleotides and conjugated nucleotides.
- 85. (New) The DNA polymerase of claim 10 wherein said mutation in Region II is selected from the group consisting of: a leucine to histidine mutation at a site corresponding to L408 of SEQ ID NO: 2; a leucine to phenylalanine mutation at a site corresponding to L408 of SEQ ID NO: 2; a proline to leucine mutation at a site corresponding to P410 of SEQ ID NO: 2.
- 86. (New) The DNA polymerase of claim 10, said polymerase further comprising an alanine to threonine mutation at a site corresponding to A485 of SEQ ID NO: 2.
- 87. (New) The DNA polymerase of claim 10 that has reduced discrimination against a non-conventional nucleotide selected from the group consisting of: dideoxynucleotides, ribonucleotides and conjugated nucleotides.

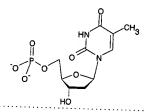
88. An isolated recombinant JDF-3 DNA polymerase that comprises an alanine to threonine mutation at a site corresponding to A485 of SEQ ID NO: 2.



CAL LIST OF COMPOUNDS

PRODUCT			USS	PRODUCT
	THYMIDINE 3'-MONOPHOSPHA	TE		THOMBER
T 8018 	Ammonium Salt Enzymatically prepared [40705-89-3] C ₁₀ H ₁₅ N ₂ O ₈ P FW 322.2 (for free acid)	25 mg 100 mg	50.50 141.20	
T 1008	Disodium Salt Chemically prepared [108320-91-8] CnoH ₁₃ N ₂ O ₈ PNa ₂ FW 366.2	25 mg 100 mg	37.80 117.90	T 2807
T 3512 -⊙€	Sodium Salt Enzymatically prepared [108320-91-8] C ₁₀ H ₁₅ N ₂ O ₈ P FW 322.2 (for free acid)	25 mg 100 mg	50.50 141.20	T 6632

THYMIDINE 5'-MONOPHOSPHATE (Thymidylic acid; TMP; dTMP)



T 9758 	Free Acid 98-100% [365-07-1] C ₁₀ H ₁₅ N ₂ O ₈ P FW 322.2	50 mg 250 mg 1 g	14.60 51.70 142.30
T 7004 ©	Disodium Salt Minimum 99% Also available as part of a kit. See: Standards and Controls Sec- tion Page 2155 [33430-62-5] C ₁₀ H ₁₃ N ₂ O ₈ PNa ₂ FV	100 mg 250 mg 1 g	12.40 24.20 66.30

THYMIDINE MONOPHOSPHATE, CYCLIC See: Thymidine 3':5'-Cyclic Monophosphate Page 956

-9°C	HYMIDINE 5'-MONOPHOSPHATE p-NITROPHENYL ESTER Sodium Salt Approx. 98%	100 mg 250 mg	29.30 80.80 177.90	ı
	A sensitive chromogenic substrate	for venom	i	

phosphodiesterase. It is not hydrolyzed by bovine spleen phosphodiesterase. [98179-10-3] C₁₆H₁₇N₅O₁₀PNa FW 465.3

4-Morpholine-N,N'-dicyclohexylcarboxamidine Salt Approx. 98% [7777778-7] C ₁₄ H ₂₂ N ₃ O ₈ P • C ₁₇ H ₃₁ N ₃ O FW 684.8 R: 20/21/22-36/37/38 S: 26-36	7 9625 -ञ्ड	WORPHOLIDATE	l g	83.20
R: 20/21/22-36/37/38 S: 26-36		Approx. 98%		
7.700 0.20-30		R: 20/21/22-36/37/38 S: 26-36	FW	684.8

THYMIDINE PHOSPHATESSee: Thymidine Mono-, Di-, or Triphosphate Page 956

- U	s
THYMIDINE PHOSPHORYLASE (Thymidine, orthophosphate deoxynbosyltransferase EC 2.4.2.4) Activity: Minimum 500 units per ml. Unit Definition: One unit will convert 1.0 μmole ead of thymidine and phosphate to thymine and 2-deoxyribose 1-phosphate per min at pH 7.4 at 25°C. [9030-23-3]	

1	1.000 20 07		
T 2807 2-8°C	From Escherichia coli; Recombinant Solution in 0.5 M potassium phosphate containing 2 mM	500 units 1,000 units 5,000 units	35.40 58.90 196.10
	uracii, 0.02% sodium azide albumin.	and bovine s	erum
T 6632	From Escherichia coli	5,000 units	254.90

-	1 6632 2-8°C	From Escherichia coli Solution in 0.5 M potassium phosphate containing 2 mM sodium azide and bovine ser Aseptically filled	uracil, 0.02% um albumin.	
1	T 0251	HYMIDINE 5'-TRIPHOSPHATE	10 mg	23.40

	T 0251	THYMIDINE 5'-TRIPHOSPHATE (dTTP; TTP)	10 mg	23.40
	-oc	Sodium Salt	25 mg	43.00
i	•	Approx. 97%	50 mg	71.30
j		[18/22/22] 0 11 11 0		
-		[18423-43-3] C ₁₀ H ₁₇ N ₂ O ₁₄ P ₃ FW 482.2 (for free acid)		in dry ice
1		R: 23/24/25-36/37/38 S: 45-26	5-36-22	
- 1		TINAMENT TO THE PARTY OF THE PA		

TINGSIDE
THYMIDINE-5'-TRIPHOSPHATE-[METHYL-3H]
See: Radiochemicals Page 2148
THYMIDYLIC ACID

1		See: Thymidine 5'-Monophosphate	Page 957	
	, ,	HYMIDYLYL(3'→5')-2'-DEOXY- ADENOSINE Ammonium Satt	10 mg	83.7

1	-o-c	Ammonium S	alt	
		[61845-39-4] acid)	C ₂₀ H ₂₅ N ₇ O ₁₀ P	FW 554.4 (for free
ı		D. C		

113508	CYTIDINE	5 mg	51.9
-0°C	Ammonium Salt		
	[61845-38-3] C ₁₉ H ₂₆ N ₅ O ₁₁ P • NH ₃	FW 548.4	

T 3883	"HYMIDYLYL(3'- GUANOSINE Ammonium Sa Approx, 95%	→5′}-2′-DEOXY- lt	10 mg	229.50
		C20H26N7O11P • NH3	FW 588	8.5

_			
T 8880 ⊡©©	HYMIDYLYL(3'→5')THYMIDINE (TpT) Ammonium Sait	5 mg 25 mg	56.20 196.30

	[1969-54-6]	C20H27N4O12P • NH3	FW 563.5	
T 6633	(3'→5')-2'-DE	'→5')THYMIDYLYL-	1 mg	41.20
	Disodium Sal [108347-86-0	t)] C ₂₉ H ₃₇ N ₇ O ₁₈ P ₂ Na ₂	FW 879.6	

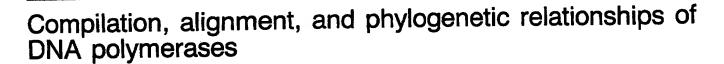
_		0/ 3.0	
T 0376	HYMINE (2,4-Dihydroxy-5-methylpyrimidine; 5-Methyluraci) Minimum 99% Also available as part of a kit. See: Standards and Controls Section	25 g 100 g	8.40 12.80 24.90 73.50
	occ. Ottaindards and Controls Section	1 Paga 216	2

See: Standards and Controls Section Page 2162 See also: Tissue Culture Media and Reagents Page 1776 and Page 1871 [65714] C ₅ H ₆ N ₂ O ₂ FW 126.1
THYMINE-2-14C

See: Radiochemicals Section Page 2148
THYMINE-METHYL-3H
See: Radiochemicals Section Page 2148

	Tradiochemicals Secur	on rage 2148	
T 3766 2-80	HYMINE 1-β-o-ARABINO- FURANOSIDE (1-β-o-Arabinofuranosylthym Minimum 99% [605-23-2] C ₁₀ H ₁₄ N ₂ O ₆ FI		173.00 576.40





Dan K.Braithwaite and Junetsu Ito*

Department of Microbiology and Immunology, College of Medicine, Arizona Health Sciences Center, Tucson, AZ 85724, USA

Received January 14, 1993; Accepted January 22, 1993

INTRODUCTION

This is an update of an earlier compilation and alignment of DNA polymerase sequences (Ito and Braithwaite, 1991). As in the previous compilation, we attempted to compile complete sequences, to facilitate the identification of conserved and viable regions of the DNA polymerases (1). This update includes, for the first time, three DNA polymerase sequences from Archaea (2); two new members of the Family A DNA polymerases; and 19 new members of the Family B DNA polymerases. In addition, we included nucleases that have related amino acid sequences to E.coli DNA polymerase I, and the sequence of E.coli DNA polymerase III (e-subunit) was aligned to Family C due to its homology to Bacillus subtilis DNA polymerase III.

As in the previous compilation (1), Family A DNA polymerases are named for their homology to the product of the polA gene specifying E.coli DNA polymerase I; Family B DNA polymerases are named for their homology to the product of the polB gene encoding E.coli DNA polymerase II; and Family C DNA polymerases are named for their homology to the product of the polC encoding E.coli DNA polymerase III alpha subunit.

Table 1 summarizes the molecular weights and isoelectric points of each DNA polymerase and nuclease. Table 1 also serves as a reference guide to the sequences shown in Figures 1A, 1B, and 1C. Since no new sequences were published for the Family X DNA polymerases (β -like), we have excluded them from this compilation.

...

SEQUENCE ALIGNMENT

The multiple alignments of the amino acid sequences for this update were performed in most cases by merely adding on to our original alignments (1) where possible. Due to the large number of sequences added to the alignment for Family B we have changed the original alignment in some areas between obvious blocks of conserved sequences. The newer sequences were added by aligning each to the closest related sequence already aligned, or in many cases to the closest related group f sequences already aligned. A more recent addition to the UWGCG (University of Wisconsin Genetic Computer Group) program package, PILEUP, a multiple alignment program, was used extensively to try and locate significant homology in groups

of closely related sequences. These newly formed groups of highly related sequences were then regapped to conform with the entire alignment based upon the previous alignment of those sequences in the new group from the original alignment. As in the previous paper, all the final adjustments had to be made by eye and, as stated above, in Family B the added sequences led to some improvements to the original alignment that became evident to the eye when they were being combined with the entire alignment by hand.

GENERATION OF PHYLOGENETIC TREES FOR THE DNA POLYMERASE DOMAINS

Using Felsenstein's PHYLIP program package (71), specifically the programs named in the outline below, we generated phylogenetic trees for the 9 Family A DNA polymerases (Figures 2A and 2B) and for the 47 Family B DNA polymerases (Figures 3A and 3B). The trees for Family A were created from the alignment in Figure 1A using the most conserved regions found at the following positions: 798 to 814, 877 to 998, 1047 to 1090, 1104 to 1123, 1131 to 1158, 1175 to 1206, 1236 to 1251, 1284 to 1305, 1322 to 1340, and 1365 to 1379. These conserved regions were recombined and 100 bootstrap samples were generated using SEQBOOT program. Using the DNADIST program, we turned the samples into distance matrices using the Kimura-2 parameter method. The resulting matrices were then input to the NEIGHBOR program using the UPGMA method to produce approximately 100 trees. Finally those trees were reduced to a single tree using the CONSENSE program. This final tree was then plotted for publication using two different methods. The trees in Figures 2A and 3A were created by the DRAWGRAM program setup to produce a phenogram type tree and the trees in Figures 2B and 3B were created by the DRAWTREE program. The trees for Family B were created from the alignment in Figure 1B, according to the same procedure, using the most conserved regions found at the following positions: 1407 to 1760, 1885 to 1901, 1956 to 1990, 2081 to 2100, 2181 to 2210, and 2280 to 2320. The Family B DNA polymerases can be subdivided into two subfamilies, the protein-primed DNA polymerase subfamily and the RNA-primed DNA polymerase subfamily.

^{*} To whom correspondence should be addressed

Mucierc	ACKES ILL	, c, c,	•		_		۰	928883833	8	SEE SEE SEE
m-60-60	SECRE		223.23 8	1222 SEE	8	553555 5 03	s	••••		_
8-3333 :	PACE DECTED DECT	AOD AAUPGÉ Vélitio	SEPTAS	600 645 645 645 645	8	FENIENCY TYCHEOPLY TYCHEOPLS TEEVEROLS TEEVEROLS TYCHEOPLS TYCHEOPLS TYCHEOPLS		VARIED VARIED VARIED VARIED VARIATION VARIED VARIATION VARIED VARIATION VARI		1200 1200
TEVPOT	265 265 265	ALO AN				SENCOL LAVER I		25222225 25222225	TI CONT.	747605742 747605742 7447474
		A GAKG	9844			200 mg		56555555 56555555 56555555		
اننين		VEAGO		- 보면 : 128 - 128		- 100 A 10	-	EGENERAL SECTION OF THE PROPERTY OF THE PROPER	Cerss	THE PARTY OF THE P
SEPTEMBELLO PFIRELLO ALIKELO A	200 200	S CONTAG		S75 LEGPKALE SERGTHGL	K	TETTE TETTE		PERTATAL PER	#tt-tckLss	K 7372775EE
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		350 375 375 350 375 350 375 350 375 350 375 350 375 350 375 375 375 375 375 375 375 375 375 375	74757 74762	1007-12-1		GUES TAGOTI AGGIOTTA (GI		PT CANTER STATE OF THE STATE OF		EGAAKTEEPTVE
8800	255 255	וופרט	CCCYELCON CCCTYELCON TREGORALI	25		HEACK AGERA AGERA LARBSE				TE EOW
4 444	AAAAA	PAAEE	E CONTRACTOR		g ::: ≱ g ::: ::			2-555×395		
2355	PERSE PERS PERS	TEV00	PPILTA AGAZGI GO TILLA DE EL ENTRE EL EL ENTRE EL ENTRE EL ENTRE EL ENTRE EL ENTRE E	PRESCENT AND	- Hoskar	ANT LACE ENGIPPO TALLEELY ANT LACES TOOST VALLEELY				
ESTATES TATES	222		T TOUR TENT	250 FIRATE	<u> </u>	g. 144 S		20 - 12 - 12 - 12 - 12 - 12 - 12 - 12 -		1150 120 120 120 120 120 120 120 120 120 12
150 DELF.EHYK TEMY.APYK MEAY.GGYK	27.5	Tovelle	PONT SELECTION OF THE S	550 ELECT CONTROL OF THE CONTROL OF		500A7		2222222 2222222		THE STATE OF THE S
	DEVIL	EATIK	71. 15. 15. 15. 15. 15. 15. 15. 15. 15. 1	AVSC. PASC. PEREN VELER VELER VELER VELER		Alac Curson		WILEYRGIAKI KILDYRELIKU RILOYRELIKU RILOYRELIKU RILOYRELIKU RILOYRE ROBENTORK KAREMULIKU	5	PSAKA PRAMICA PONACI PASCICA PERCECA P
WFPA	2257	NI BL	PROCESSOR AND	2007 2007 2007 2007 2007 2007 2007 2007		VEDRG TEMAA DANDY				
\$458 \$458	SACA STATA	GHKEV	YEREST S	25.55.55 25.55.55 25.55.55 25.55.55 25.55.55 25.55.55 25.55.55 25.55.55 25.55	MV. AY	757		~ XXXXXX X		
125 INOTITIMA. LENEDSHILL LENEDSHILL LENEDSHILL LENEDSHILL	Tage	X-X	CENTRAL SECTION OF THE SECTION OF TH	SECTION SECTIO	2	727 Truses Livers		SSOCIAL SECTION SECTIO		112 ARREDIT ARREST ARRE
3333	AKST KEST KEST KEST KEST KEST KEST KEST K	RGAKTI	POEIX ACTED	PASSES EST		SLYYD		22222222 23222222 232232222		PEVEG PEVEG SPASA
PER	200	AGLSF	APRING PT	SALEA SALEA				STANDER OF THE PROPERTY OF THE		######################################
PTGAY PYCAY PYCAY	ASSY		TOTAL SOCIAL	ALE PAIRE EVANT	\$			20204442 20204442		000000000000000000000000000000000000000
9993			SICKET TENESTY WENT	200 200 200 200 200 200 200 200 200 200		TOO SHARE PEREKY SWARN GGEVI SWARN GGEVI SWARN GGEVI LEEL LSOKHYP IKHR SCHYPP IKHR SCHYPP IKHR SPYFIKKE		8. : 44		TOO TOWNERS FORMERS FO
§. ≾≤97	## ## ## \$	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	TANGE OF THE PERSON OF THE PER	75		TANGET IN THE PROPERTY OF THE		ਰਕਕਕਕਰਕ ਉਨ੍ਹਾਂ ਵਾਲੇ ਜ਼ਿਲ੍ਹਾ	Ş	2 3
	E SE	A STAN	CATE CONTRACTOR	9 65		SECTION AND AND AND AND AND AND AND AND AND AN	8	535555555 535555555		ZZZZZ
999		223 223 223		Polska				TO TO TO THE TO	EGVET	SECOND LIAN SCORE LIAN LAGE FAIL LYGOS FANKS
RAYNA RAFFA RAFFA FFA	STALV	222	ACKILL STATE	A 1100	a a	ALECTO MEETING OF THE PROPERTY		MANAGE TO SEE THE SEE	\$	AAH SA AAH SA AAH SA AAH SA AAH SA AAH SA AAH SA AAH SA AAH SA AAH SA
ESSS 2	252 252 252	25.00 660 73.00 70 70 70 70 70 70 70 70 70 70 70 70 7		475 475 WAEGLDF	55	2. S.		£	5	
2555	999		YEALER FYRKER FYRKER FLEATING	A 475 A 77 A 77 A 77 A 77 A 840E	T CCS			PT PY	3	65555555555555555555555555555555555555
9-19-19-19-19-19-19-19-19-19-19-19-19-19	EXCO EXC EXC EXC EXC EXC EXC EXC EXC EXC EXC			SFALEP SOALNE SISHOE VIYOYS	12 12 13	SECTION AND A SE			3	SYLLYS SERVING
₹ 153 5 153		855 855		# 52 GBS	<u> </u>	23500 23500			d .	\$4.65\$************************************
	55	ACE ACE	SKICAR	PADEN FYDEN YCRDG	VSD T	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		0.		MARKAFIL KRAFF
۶,	E8	2	AKGOG NEGOG	450 FETDS: FETDS: FETDS: FETS:		650 WOLEAN DEENAN WELGEN		6		25 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
	Age .	CKGVA KGVA KT HP	ALL BEER IAILUPEUTITPAKA Nasaldilee Andropologiseka 1 paggileettä	YEAFD YSCPO YS YSCPO YSCPO YSCPO YSCPO YSCPO YSCPO YSCPO YSCPO YSCPO YSCPO YSC	33	- PAGE STATE	. YS 101	AGE FACE FACE FACE FACE FACE FACE FACE FAC	LEGFERING PROTUBER (1) 115.81 (1) (1) (1) (1) (1) (1) (1) (1) (1) (1	2000 A
	9		2 35	PART OF THE PROPERTY OF THE PR	44 44 64 64 64 64 64 64 64 64 64 64 64 6	HEADER HE	ĝ.			
				WALLAND I	APVE.				TVEPK	BAVEN
		200 200 200 200 200 200 200 200 200 200			20 20 20 20 20 20 20 20 20 20 20 20 20 2	\$255555 \$255555 \$255555 \$255555 \$25555 \$25555 \$2	AKS	S TONGE S		8
χ,		2 88 55			8) N			
	50.00	\$-0.50 0.00 0.00 0.00 0.00 0.00 0.00 0.00		VISTO KEGAFI	AELCA	25 95 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	WSAL		F	
	No.	E C		EVIAT EKIAK	TANAL ENAME ENAM ENAM	F 55	700A	ATOVR ATOVR PKYOL PKYOL LDOE	ACCON	9
	MSADLYT I PROVADI 1 GOT 10 SLEKEROŠLKROLINEROETVAELEEKLIKOT LOLICO PRESIDENTI	OZ ZZ ZZ WINNIN WELL WAS AND W	PYTALLANDER EGGEWILF KOOD VOLLEER IN LIPEGOLI IPAAA Y Maskidileerin oo	Danyi, Vezi e ilegiri i i i saggaria 1 S-0-0 - (25		Cop	E I E I ED TÖGSALINYSALIKSLICOVAKFIKIKI DLUKTURDFF ASTUK	PAL SATERBRINKTO DAYL HAISESELT TRUNCTE CLAME LAGE FALLOUS LIGHTWATER ILL STRUCKTE LAGE FALLOUS LIGHTWATER ILL BROWNER IN BETT TO SET THE LAGE FALLOUS LIGHTWATER ILL STRUCKTE LAGE FALLOUS LIGHTWATER ILL STRUCKTE LAGE FOR LIGHTWATER ILL STRUCKTE FOR LIGHTWATER ILL STRUCKTE FOR LIGHTWATER ILL STRUCKTE FOR		1050 CAR ROALINE ROAL
· · · · · · · · · · · · · · · · · · ·	:::::::::::::::::::::::::::::::::::::				:32:::: '35≥ ::::			2 NGCANG : Testing		\$35555555 S
										E-882005-13
888	7. 7. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2.	### 100 #### 100 #######################		74.57 74.57 8.88 8.88 8.88 8.88 8.88 8.88 8.88 8	200 - 12 - 12 - 12 - 12 - 12 - 12 - 12 -	8888	77. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7. 7	8888	Spirit Strain St	#####################################
ح شوب ح	.			لتدر المتالية	Little O	mini⊑ ,	177	men-		#### P P P

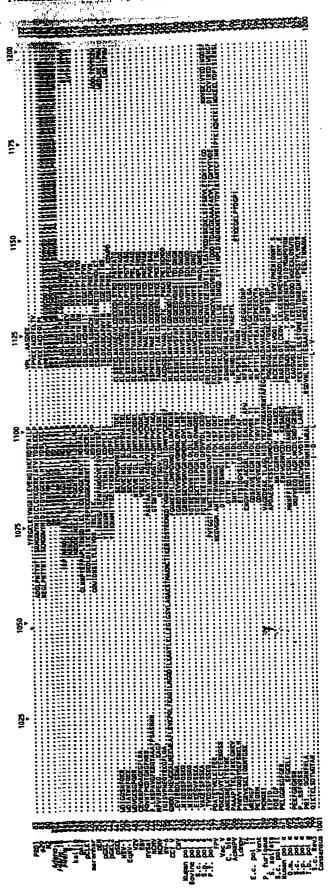
	~	<u> </u>
25.000 (20.000) (20.0	90 ·	10000000000000000000000000000000000000
1373 LOVALLYNGGERONGAR LOVALLYNGGERONGAR LOVALLYNGGERONGAR GRANDAR COLLOGERONGAR GRANDAR COLLOGERONGAR GRANDAR COLLOGERONGAR GRANDAR COLLOGERONGAR GRANDAR COLLOGERONGAR GRANDAR COLLOGERONGAR APPERIOR CERROR COLLOGERONGAR APPERIOR COLLOGERONGAR APPERI	Řτ	ANI PYTE LIPSOLOSION SANISTER CONTROL SA
1330 TEST VENEZIONAMENTES PENEZIONAMENTES PENEZIONAMENTE PENEZIONAMENTE PENEZIONAMENTE PENEZIONAMENTE PENEZION	- -	REGION DESCRIPTION OF THE PROPERTY OF THE PROP
1325 1325	χ₃-	TOWN GRADOUTS AS \$19 (1) AND ASSESSED AS \$15 (1) AND AS \$15 (1)
1300 1300	9≱	A CHARGOON WOOG GIVED ON THE CHARGO ON THE C
1273 CARETAN INST CARETAN IN	βCF	WEALP RESOUCESHEEL CARE
1230 SECTOR 1 (1991)	α•	Herves sould assessed and the control of the contro
TATE OF THE CONTRIBUTE OF THE	æ* ∞*	ASSELIATIONS CONTRACTOR CONTRACTO
8000 000 000 000 000 000 000 000 000 00	~ {************************************	
Signer	g a	<u> </u>

TO MONTH SOURCE THE	***	chāratigasaningasesarinassaras aceasaigatara
8	224-224 1-24-22-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2-	
A CONTRACTOR CONTRACTO	MANAGE TO THE PROPERTY OF THE	16.11. 11. GENGWYNER ITTE GENEGGAAAGETE BEGENFER AND STATE OF THE PROPERTY OF
PPP PPP PPP PPP PPP PPP PPP PPP PPP PP	A TONERA	
WELLE OF THE PROPERTY OF THE P	25.00	
TT. THE LEGISLE OF TH	TAN TO THE TANK THE T	
PERSONAL PROPERTY OF THE PROPE	Para Para Para Para Para Para Para Para	######################################
A SCARCE CONTRACTOR CO	VOVETOR VOLKE ESGARE ESCARE ES	
THE STATE OF THE S	EST CONTRACTOR CONTRAC	
350 ALLANDA	NAME OF THE PARTY	
AND THE PROPERTY OF THE PROPER	2012 2012 2012 2012 2012 2012 2012 2012	THE PARTY OF THE P
7 7 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	PETER	
WALES	KERALDE KERALD	AND THE PROPERTY OF THE PROPER
22 22 22 22 22 22 23 24 24 24 25 25 25 25 25 25 25 25 25 25 25 25 25	25 - 25 - 25 - 25 - 25 - 25 - 25 - 25 -	
SECTION SECTIO	DAAA T SAGO SAGO T SAGO T SAGO	
Session Control of Con	ANANA MANANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANA MANANANA MANANANA MANANANA MANANANA MANANANAN	DESCRIPTION OF THE PROPERTY OF
300 300 300 300 300 300 300 300 300 300	######################################	: : : : : : : :::::::::::::::::::::::::
300 200 200 200 200 200 200 200	1.00 P. 2.00 P	00
2000 N. 100 N. 1	A CERT	STATE
	LO MOR	######################################
# 100 man	KING KAN	THE CONTRIBUTION OF THE CO
272 273 273 273 273 273 273 273 273 273	TOTAL SELECTION OF THE	
25 - 25 - 25 - 25 - 25 - 25 - 25 - 25 -	A SA	
	Se resident	
250 CENTRAL CANADA CANA	WANTER WA	S. S
250		4-50 (1.00)
	3 600 3 600 3 600	221 2010 2010 2010 2010 2010 2010 2010
KI-LITER TO THE		
= = = = = = = = = = = = = = = = = = =	NAI III SO	TAN TO THE TOTAL
POSZ SZ S	AND STATE OF THE S	KAGANTALSHAELLING VYPEIG. I NICHELLING WAS THE STANDARD
	######################################	### ### ##############################
454 E E E E E	Estation of the state of the st	▲ 경우 및 경우 및 등록

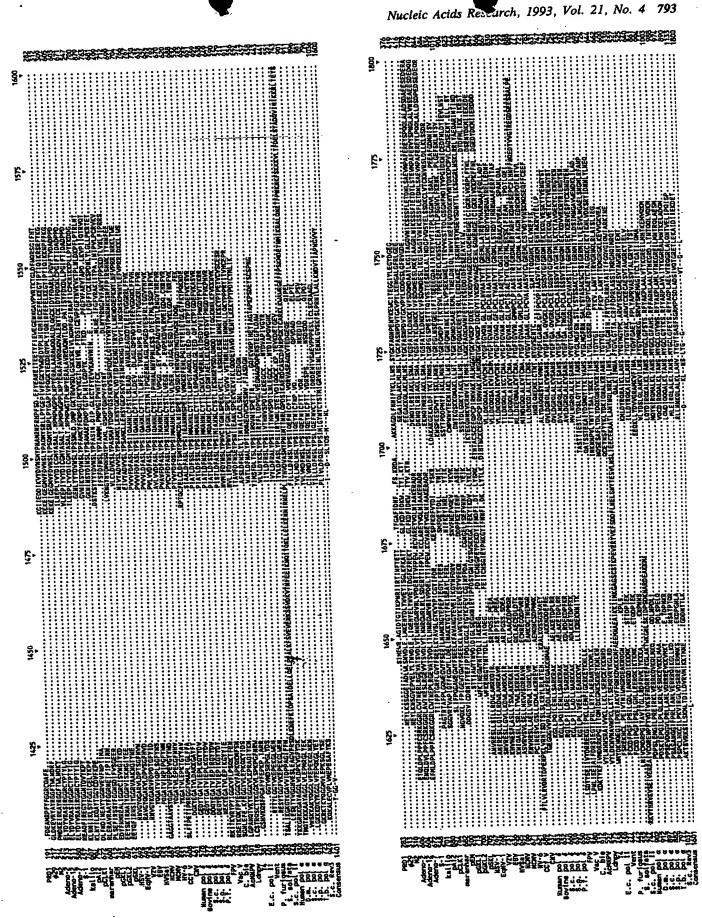
	\$1750454545555545555555555555555555555555	
g• :	STATE OF STA	
000 TT 00T 5T 00T 5T0	SONO CONTENT AND THE CONTENT A	
5°	Lining I K 1905 185 50/5150 IIII Lining	-
12.	DESCRIPTION OF THE PROPERTY OF	
904	STATESTICE CANADATE IN THE STATESTICE CONTROLLED TO THE STATESTICE CONTROL	-
£.	GOOGNAMICATURE LIGHT IN CONTROLL CONTRO	
9	Spondarsav 1 Control of the control	
% •	CONTRACTOR	
**		_
	24 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Consensus

1040	r erectively	######################################	initial distribution	Manageria
8- ::			54455555555555555555555555555555555555	
			ZEELAKE KAKE	
		2555555E		2421 488255
		2000 N 1000 N 10	E CERTA	2655555555 27555555555555555555555555555
•		2222200		.\$4 2 00
Š				
٧. ا	2244420045			£######
	149984		223333E5E55	
5		**********		
Š		20000000000000000000000000000000000000		
3				
¥				22222222222222222222222222222222222222
EP		28 : 36 2 2 2 dd		
2.2	SIGGPRELLENGE CONTROL OF THE CONTROL	:25255588 <u>25</u> 2		
				######################################
Ę			72322222222222222222222222222222222222	**************************************
179	캶톲딦궦찞퍞퍝찞푶			
. 5			22222222222222222222222222222222222222	
9	25.525.52 25.525.52 25.535.535.535.535.535.535.535.535.535.5		TANK STATES	ANTHARA THE THE PROPERTY OF TH
Ð	POTO INTO THE PROPERTY OF THE	THE CONTROL OF THE CO	8	MACKADARWAGE SAME AND STATE OF THE STATE OF
8.	JSG 1870 HL OCONISS LOVY A THE HEAVING THE SERVING T	LIMBOTEL MSEEN IN	KTGKKKTVVTNICOGRAMMEN PVERLALDA	
	VSSTVASIPHLEGUALS WATFLEGUAPHISCUALS WATFLEGUAPHISCUALS WATFLEGUAPHISCUALS WATERLAND TRANSPORTER	25 25 25 25 25 25 25 25 25 25 25 25 25 2	2	
	A A P P P		Ē	
	222	3	KIÇDIKKETWITH I KOĞBAT	9
				3
		ggg	\$	
8.		T. S. A. C. W. B. C. S.		
8∙	255		Š	
	244		Ebagaga	
	PETRI IPECANICAORITEORICACIANO PARTICI IPECANICARRAMINA CALMADI PARTICI IPECANICARRAMINA CALMADI PARTICI IPECANICARRAMINA CALMADI PARTICI IPECANICARRAMINA CALMADIA	Ling 16, 2500 (PESILIFE LING)	\$255005 \$2555555555555555555555555555555	ELEMENTOC.
		######################################		\$
	200			3
	500		¥ = = = = = = = = = = = = = = = = = = =	₹9
Ę.		\$5		
		111 11 11 11 11 11 11 11 11 11 11 11 11	TOLPOSPIVATIONELEE TEINTO I FORESTILINADE TEINTO I FORESTILINADE TEINTO I FORESTI	\$\$500 \$400 \$400 \$400 \$400 \$400 \$400 \$400
	262			SA SAME
	202	ZZ		
			\$2555555555555555555555555555555555555	
			20000005555 <u>5</u>	
_	2325556 2325556			PACHER POACH
₹.	2228262	25252525252525252525252525252525252525		24444 24444 24444 24444

		25 222222	7877557	######################################
			777 FANT 27 - 72 - 72 - 72 - 72 - 72 - 72 - 72	\$255 X FEATER
	2444 244 244 244 244 244 244 244 244 24	\$	255	\$2555555555555555555555555555555555555
	2 20 E	PU: 252. Edu	28-2-2-6-2-6	2 - E - E - E - E - E - E - E - E - E -
χ.		2027 2000 000 000 000 000 000 000 000 00	~ ************************************	2072 - Yeres
10	*********	20000000000000000000000000000000000000	0 - 2 3 3 3 3 5 8 5 8 5 8 5 8 5 8 5 8 5 8 5 8	
	5445553445 FILTER PROPERTY OF	55955555555555 	229222222	26855555555
	SAREESEE		24422222222222222222222222222222222222	13 53 53 54 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
	产業を自己の子子ご		¥ ŽŽŽŽŽ ĘŽ <u>S</u>	ZEZ-28662 ZZ
•	24 500E-5		522 222 mor	# 2500 H
	\$:00050E			
	* : :2222===	1 102 10 11 11 11		THE STATE OF THE S
	COLOR DIGITAL	The second second		
	244 - 12 - 24 - 24 - 24 - 24 - 24 - 24 -		\$25 88888 %	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
	565 3	서는 48 대 화 중==	A RENEE P	2
	3-	4	Tonno Egiqei	
			-3	



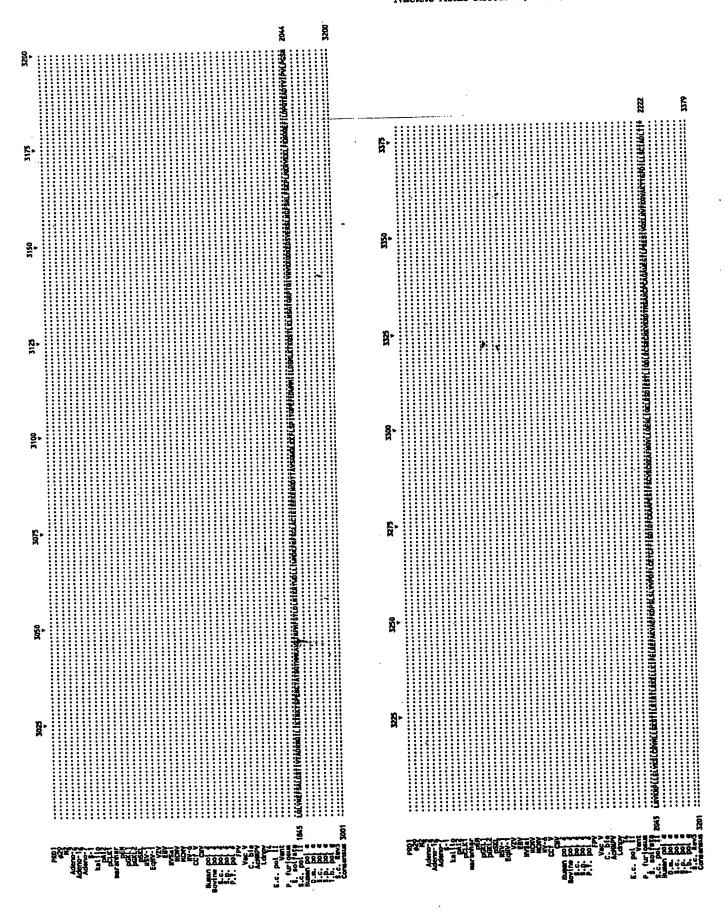
8		23E 25E 25E 25E	Senencial Control
8+ :	4444464110022455		¥ : : : : : : : : : : : : : : : : : : :
			I LAKSSALI
-	MAN (1995) (1995	HIKAL BAGIWA TOT I	
	######################################	SKAR SKAR SKER SKER SKER SKER SKER SKER SKER SKE	
		A	
Ë,		5	3
	MANALEEANASEAS PARAMETER PRODUCE ERECTISSMILLE POCCECCL PROFILE OF STATE OF	SSACVALDACTVANTOV ISIN	OOF GALES TO SALE OF TROPY FRATTOL SALES PERSONALIDIS CALLER PROJAMA PPARAL PROJAMA PROJAMA PPARAL PROJAMA PROJAMA PPARAL PROJ
	TOGRĖTĖS LOCZBANTPANIE SERIADSEREGOSISE EERI TOGRĖTĖS LOCZBANTPANIE SERIADSEREGOSISE EERI TOGRĖTĖS LOCZBANTPANIE SERIADSEREGOSISE EERI TOGRĖTĖS LOCZBANTPANIE SERIADSEREGOSISE EERI TOGRĖTĖS LOCZBANTSERIA SERIADSERIA TOGRĖTĖS TOGRĖTĖS LOCATORISE LOCAT	160	3
		1972	
		14. 19.	<u> </u>
38	2		a .
_	TOGRĖTĖS GAZBOTIONAME KASIN DEFERENCERIES TOGRĖTĖS TOGRĖTĖS GAZBOTIONAME KASIN DEFERENCERIES TOGRĖTĖS		CEPPHIER DEFETURE PREFETURE THE PROPERTY OF THE SAME PROPERTY OF THE SAME THE PROPERTY OF THE SAME THE SAME THE PROPERTY OF THE SAME THE SAME THE SAME THE SAME THE PROPERTY OF THE SAME THE SA
	PAARE FEYK FAVY EAAAG		
	AX: 54%: 64%: 64%: 64%: 64%: 64%: 64%: 64%: 6		2004m
	\$ 50 00 00 00 00 00 00 00 00 00 00 00 00		
χ. •	68 1 1970 11 1970 12 1970 13 1970 13 1970 13 1970 13 1970 15 1		
	20 - 10 - 10 - 10 - 10 - 10 - 10 - 10 -	\$2000E	
	TERMINE TO THE PROPERTY OF THE	LPSAYS TO ANY KEEL TOWN KEEL TOWN KEEL TOWN TOWN TO ANY TO	VI POOL BEEFFURTH TO THE POOL BEEFFURTH TO T
	VVALED VV	TOTAL DEVISE TO THE TOTAL	
	AND THE STATE OF T	SALE EALERMAN I. PENNS SELECTION PROPERTY SELECTION	
8.	:555447474747474747474747 :575447474747474747474747474747474747474		
	77875 17886 18596 18596	A STANDARY SEE	
		KALEGOO & & & & & & & & & & & & & & & & & &	STATES TO THE STATE OF THE STAT
	AND		PARTICIPATION OF THE PROPERTY
	PANEL STATES OF THE STATES OF	**************************************	POPULATION OF THE POPULATION O
1273	######################################	MARKA	20000000000000000000000000000000000000
	A STATE OF THE STA		
	A A A A A A A A A A A A A A A A A A A		VYALE VYALE VYALE TYPEL
		25000000000000000000000000000000000000	SERVICE STATES OF THE PROPERTY
<u>გ</u> '	\$5555555555555555555555555555555555555		######################################
	SECOND SE	SERECTER SERVICES	SEAST SECTION OF THE
	APPENDENCE STATES	25-28-28-28-28-28-28-28-28-28-28-28-28-28-	# # # # # # # # # # # # # # # # # # #
	CAN ESCRIPTION	ZONTING STATE	A AAAA
		<u> </u>	
23		966	
		VIII I	GE.
		200	
	THE TOWNSHIP TOWN THE	TANKORY MATERIAL PROPERTY OF THE STANKE STANKES STANKE	TOWN AND STATE OF THE STATE OF
	A CONTRACTOR OF THE CONTRACTOR		
		5-6-55- 5-6-55- 5-	
	565 11 - 11 11 11 11 11 11 11 11 11 11 11 1	XXXXX 2	132 8 528
	~ *	2 0 00-	



and the second s
10011111202-1-1-1-1-2-2-2-2-2-2-2-2-2-2-
60 9999 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
99977799999999999999999999999999999999
1950 1970
THE THE TABLE TO T
POST OF THE PROPERTY OF THE PR
NEW TO SEE THE PROPERTY OF THE
0
ANTER CONTRACT CONTRA
100 100 100 100 100 100 100 100 100 100
3 274 3 283 3 833
g- 909
TEAL AND THE TANK THE
00.00 0.00 19.00 0.00
THE RESERVED THE PROPERTY OF T
24.01.00.01.00.01.00.00.00.00.00.00.00.00.
454 8 25440 1 4942 4444 444 444 444 444 444 444 444 4

######################################
CONTROL OF THE PROPERTY OF THE
STATE OF THE BEAT
CHICAGO III THE LEAGONT LIGIDEE TO THE LEAGONT LIGIDE TO THE LEAGONT LIGIDET LIGIDE
2550 2750 2760 2770 2771 2771 2771 2771 2771 2771 277
THE CONTRACTOR WIDNES TO SEE THE CONTRACTOR WIDNES THE CONTRACTOR WIDNES THE CONTRACTOR WINNES THE CONTRACTOR
TOTO TOTO THE EAST TOTO TOTO TOTO TOTO TOTO TOTO TOTO T
2100 2100 2100 2100 2100 2100 2100 2100
A 1902 D 1 WERDAGI VEGSPO D 1 WERDAGI VEGSPO
TO THE STATE OF TH
2011 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
582 587.101.118
202 201 - 10
ZOGS ZOGO
2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3
244 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

######################################		<u> </u>	1013	<u> </u>
87	897			AAIPO CAA Spinskasect
A A A A A A A A A A A A A A A A A A A				Prost
				Scient 1
1				AACANC AACANC TO TO T
	K.			2007 F.
23. 17. 17. 17. 17. 17. 17. 17. 17. 17. 17				SECRETARIAN SECRET
				2500000 2500000000000000000000000000000
THE FAME PRODUCTARE LEADER GANGAGE TO THE FEETER OF THE FE				2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
0. 10.00 0.00 0.00 0.00 0.00 0.00 0.00	S.*			ACALETICS ACALET
2350 White Personal Base Figure 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1				THE TANK
THE CANALY TEAL COMMANDER TO THE PRINT PROPERTY OF THE CANALY TEAL COMMAND TO THE CANALY TEAL COMMAND TO THE CANALY TEAL COMMAND TO THE CANALY TEAL CANALY THE CANALY				TOWER T
TATE OF THE PROPERTY OF THE PR	iQ.∍			A TENTA
2322 TOTAL STREET TOTAL STREET				
				A BOALE VOOLE VOREEA
PARONE STATE OF THE STATE OF TH	8.			THE STATES
Solution of the state of the st				AND THE PARTY AN
ALED HIVESTON ALED H		22		2500000 2500000000000000000000000000000
		200 CO		SS VIII
	2	2000 2000 2000 2000 2000 2000 2000 200		
2.73 2.73 2.73 2.73 2.73 2.73 2.73 2.73				COCCEPTED AND A STREET OF THE
19		HEKKYR HEKKYR HEYKYK HEIAEK HEIAEK HEIAEK		RANKEL ROYSEL RO
9 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	957			AAL COORDER
AAPOGGEAAPAALPESPA CANTINGAROO AATACATATATATATATATATATATATATATATATATAT	957 957 957			TANCE CENTRAL CONTROL OF THE PROPERTY OF THE P
90-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1				35
				Trock And Processing Street Control of the S
50 100 100 100 100 100 100 100 100 100 1	10. 2			AVTITOR AVTITO
2225 2225 1 The control of the con		# # # # # # # # # # # # # # # # # # #		
######################################				74.0000 74.0000 74.0000
		######################################	POCYÁBA	PSTSTEE TTACKEN TTACKE
REFINAL TO STATE TO STATE THE STATE TO STATE THE STATE TO STATE THE STATE TH				######################################
A STANDARD OF THE PROPERTY OF	24.24.25.25.25.25.25.25.25.25.25.25.25.25.25.	######################################	등 등 등 등 등 - 조구 등 등 등 <u>-</u> 도	200000 20000 2000 2000 2000 2000 2000
200 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	2 15 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	#### #### ############################	بـ بـ م	
		₹ ૄૼ~~	_	m v 43



798	Nuc	leic Acids	Research,	1993, Vo	l. 21, No.	4			j.
	7 2	000	XXXXXX	**************************************	\$53 §	855 8 8	医腹 囊 蓋	85 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	9. T.C
90g <u>1</u> 11	AND SECTION OF THE PROPERTY OF	400) ASYGNESI IA TOMAWASSI POAV BAAKKAKI ISHI YOREAMI.	STS STREET, LUCKEY, CO. ST. STREET, LYKEY, CO. STRE	TO SEPTIME INV. INC. TO SEPTIME INC. TO SEPTIME INV. INC. TO SEPTIME INV	74-34-10: VOSAC 1-44-110: VOSAC 1-44-1	-d10101	#1.2-1	1575 STANDERETT CH. TOWN DETECTION TO THE STAND THE STAN	wyponi, united (Seoverier) [199 wyponi, united (Seoverier) [199 173 173 pol 1)(5), Thermus flavus DNA pol
82.		350 GELKVELNI ISPRSOROAVIGI GAL VEOL	SSO SSO SERVILLE AVIETNES THE INCLUSION OF THE INCLUSIO	TSO CELL MENER PARTICIONE I CALLE PARTICIONE I CALL	Selding of the selding	1150 DECKHPYTGFDKSVVV TOREGUMPYTGFDKSVVV TOREGUMPTHFOFFS	1350 PERMICUS SALEGUEROS 1350 1350		1750 LEMESTI PUR TYCHOLOGICE EM SLEMESTI PUR TYCHOLOGICE E
ឆ្	The Court of the C	325 Tradit wil Amonne i Kakt Red Safa	525 \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$	27 427 1903		11 - \$ - \$ - \$ - \$ - \$ - \$ - \$ - \$ - \$ -	* ::5:	DELGPIRA BRIGHRA BRIGHRA	A THE TOTAL SET THE SET OF THE SE
ğ.	PSSTERODIENSESTVEDYSPLICES	300 300 300 300 300 300 300	oos oos	NECTABLE TO LOUR THE TWO TO THE COART OF THE	OOG GOVER-LL'G	075	1275 1270 1270 1270 1270 1270 1270 1270 1270	14.75 - EL 6-4 12 1 EOW. 14.75 1 EOW. 14.75 1500 14.75 11 EOW. 1500 14.75 11	1700 1700 FOLD INTERNO DEARER 1785 FOLD FOLD INTERNO DEARER
R	wen the foster	\$12	£27	ING T GAIR HYPERFE GAIL TO SERVE GAIL LOS SALVES GAIL LOS SALV	G. EFRIAFIPOEE	VOR I SKIL VPPOPGAT	LOSGWONFIDE LOSGWONFIDE	S-MO EGYGE GP IEA KGWGEGP IEA KGMGEGP IEA	1-1-0-1-
	2	FENGEIOULTIMASSAMPHORE	BBINGEEARITVSCYVFOVETRELK 450	TOWNSTAND I THE PARKET OF THE	Wigitimorique (1978, 1978)	NEW TEACH THE TEACH TO THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE TEACH THE T	PRIVILAGI LOTAGE 121 JOSE 125 JOSE 121 JOSE 125	GTLLEPETGTR WAS LESS 1450	#. 7) -9. 80 1 1 2 2 2 2 2 2 2 2
,	ŊΡ	NECT SYNTREGEOTILE OF IMPROFERENCE TOLD THE LASS SMANNET FRANCE OF THE LASS	KESBEDAJOSEPLVIGNOI DALKE INTERSINOEEARI TVOGTVEDVE RELEKKERTI KESBEDAJOSEPLVIGNOI DALKE INTERSINOEEARI TVOGTVEDVE RELEKKERTI	**************************************	STANDA I VETI PRODUCTO I PROVINCIA SE L'ANNOCATAL SALIMAN. SALIMAN			DOPENATIONALIST GOVIETO I EXTIGITE PEGITINAMA LED 1450 1450 1450 1450 1450 1450 1450 1450	TATATITITALOGISTISTISTES. TATATATASISTES 1450 1425 1425 1425 1425 1420 1425 1420 1425 1420 1420 1420 1420 1420 1420 1420 1420
	•					# E #	뚎 호 젊	編章 融	第章 \$ \$\$ \$
	S	And a			Constitution of the second	AU CO.	8.1. DNA F 329 8.2. DNA F 320 Consensus 1001 E.c. DNA E 352	E.E. DIM B 1201	S.: PAN F 1290 Consersus 1401 B S.: PAN E 979 S.: PAN E 979 S.: PAN E 979 Consersus 1401 D

Cavicers purpues plasmid DNA pol. (pCLK1)(60), maranhar neurospora crassa DNA pol. (pSML)(62), Klayveromyces lacis plasmid DNA pol. (pGKL2)(64), Saccharomyces Klayveromyces (activate plasmid DNA pol. (pSKL)(63), herpes simplex type-1 DNA pol. (pSKL)(64), Saccharomyces Klayveromyces polymerase I (alpha)(S.c. pol o)(27), Schizosaccharomyces Pombe DNA polymerase I (alpha)(S.p. pol o)(28), Trypanosoma brucei DNA polymerase alpha (T.b. pol o)(30), and Socharomyces cerevisiae DNA polymerase II o-subunit (S.t. DNA E)(66), Salmonella typhimurium DNA polymerase III o-subunit (S.t. DNA E)(66), Salmonella typhimurium DNA polymerase III o-subunit (B.c. DNA O)(69,70).

polymerase Rev3 (S.c. Rev3)(38). C. Family C DNA polymerase III e-subunit (B.c. DNA O)(69,70). T4 DNA pol. (T4)(22), E. coli DNA pol. II (E.c. pol II)(17), Thermococcus literalis DNA pol. (Vent)(23), Pyrococcus furiosus DNA pol. (P. furious)(24), Sulfolobus solfmaricus DNA pol. (S. solfma)(25), Sacchuromyces cereviside DNA polymerase II (epsilon)(S.c. pol II)(37), Human DNA polymerase alpha (Human pol a)(26), Drosophila melanogaster DNA polymerase alpha (D.m. pol a)(29), Sacoharomyces cereviside DNA V)(50), Choristoneura bienuis DNA pol. (C. bie)(51), Autographa californica nuclear polyhedrosis virus DNA pol. (AcMNPV)(52), Lymantria dispar nuclear polyhedrosis virus DNA polymerase (LdNPV)(53), pol. (Adeno-12)(56), Adenovirus typo-7 DNA pol. (Adeno-7)(55), S-1 maize mitochondrial DNA pol. (S-1)(57), talilo neurospora intermedia DNA pol. (Ralilo)(58), Ascobolus immersus plasmid DNA pol. (pAIZ)(59), Adenovirus typo-7 DNA pol. (Adeno-7)(55), S-1 maize mitochondrial DNA pol. (S-1)(57), talilo neurospora intermedia DNA pol. (Ralilo)(58), Ascobolus immersus plasmid DNA pol. (pAIZ)(59),

Table 1. The main families and subclassifications of DNA polymerases. The number of amino acids, molecular weight and isoelectric point of each DNA polymerase or nuclease were computed using the Proteinsort program from the UWGCG. Those Family B DNA polymerases marked with a star (*) are protein-primed DNA polymerases. The sequence marked with (+) pEM (62) is not a complete sequence missing some unknown number of amino acids on the N-terminus and so the numbers is the table only represent this short sequence.

	Amino acids (No.)	Mol. Wt.	Isoelectric pt.	Reference
Family A DNA polymerases				(2)
Bacterial DNA polymerases	928	103,117	5.37	(3)
a) E.coti DNA polymerase I	877	99,078	4.78	(4)
b) Streptococcus pneumoniae DNA polymerase I	832	93,909	6.38	(5)
c) Thermus aquaticus DNA polymerase I	831	93,783	6.00	(6)
d) Thermus flavus DNA polymerase I	651	• •		
Bacter sphage DNA polymerases	020	94,410	6.19	(7)
a) T5 DNA polymerase	829	79,691	6.45	(8)
b) T7 DNA polymerase	704	106,808	5.34	(9)
c) Spot Oat A malamana	924		8.50	(10)
c) Spot DNA polymerase	648	72,561	0.50	
d) Spo2 DNA polymerase		440 450	9.23	(11,12,13)
Mitochendriai DNA polymerase	1254	143,479	7.20	(,,
Yeast mitochondrial DNA polymerase (MIP1)	e I		0.00	(14,15)
5' to 3' Exonucleases with homologous sequences to E.coli DNA polymeras	305	35,558	9.00	
a) T4 KNase H (gp 33.2)	291 .	33,448	5.12	(7,16)
b) T5 Exorniclease (gp D15)	348	40,126	4.54	(4,8)
c) T7 Exonuclease (gp 6)	J			
		- 1		
Family 8 DNA polymerases		•		/1 ** \
Bacterial DNA polymerase	783	90,020	6.85	(17)
E.coli DNA polymerase II				(10 10)
Bacterisphage DNA polymerases	553	63,336	6.68	(18,19)
a) PRDI DNA polymerase*	575	66,714	8.83	(20)
b) \$\phi 29 DNA polymerase*	572	66,423	7. 69	(21)
c) M2 I)NA polymerase*	898	103,609	6.20	(22)
d) T4 JiNA polymerase	070	,		
Archaetacterial DNA polymerases	474	89,913	8.29	(23)
a) Thermococcus litoralis DNA polymerase (Vent)	774	90,112	7.92	(24)
b) Pyraniccus furiosus DNA polymerase	775		9.72	(25)
c) Sulfation of the Control of the C	882	101,332	7.12	- ()
c) Sulfutibus solfataricus DNA polymerase				
Eukaryotic Cell DNA polymerases				(76)
(1) DNA polymerase alpha	1,462	1 65,859	5.71	(26)
a) Human DNA polymerase (alpha)	1,468	166,776	6.14	(27)
b) S. cerevisiae DNA polymerase I (alpha)	1,405	159,348	6.85	(28)
c) S.pombe DNA polymerase I (alpha)	1,505	171,167	8.22	(29)
d) Drusophila melanogaster DNA polymerase (alpha)	1,339	151,611	6.39	(30)
e) Trypanosoma brucei DNA polymerase (alpha)	1,337			
(2) DNA polymerase delta	1 107	123,634	6.94	(31,32)
a) Human DNA polymerase (delta)	1,107	123,707	7.52	(33)
b) Briving DNA polymerase (delta)	1,106	· <u>-</u>	7.96	(34)
c) S. rerevisiae DNA polymerase III (delta)	1,097	124,618	7.63	(35)
d) S. pombe DNA polymerase III (delta)	1,084	123,211	8.76	(36)
e) Phiamodium falciparum DNA polymerase (delta)	1,094	126,883	0.70	(50)
			c 00	(37)
(3) UNA polymerase epsilon	2,222	255,669	6.92	(37)
S. cerevialue DNA polymerase II (epsilon)	•		0.55	(20)
(4) Other cukaryotic DNA polymerases	1,504	172,956	8.86	(38)
S. cereviniae DNA polymerase Rev3				/am
. Viral i)NA polymerases	1,235	136,547	7.35	(39)
a) Hernes Simplex virus type 1 DNA polymerase	1,220	135,955	6.67	(40)
b) Equine herpes virus type 1 DNA polymerase		134,047	7.80	(41)
c) Varicella-Zoster virus DNA polymerase	1,194	113,417	7.38	(42)
d) Epstein-Barr virus DNA polymerase	1,015	113,934	7.31	(43)
e) Herpesvirus saimiri DNA polymerase	1,009		7.25	(44)
f) Human cytomegalovirus DNA polymerase	1,242	137,101		(45)
g) Murine cytomegalovirus DNA polymerase	1,097	123,573	6.68	(46)
b) Hunter beauty with the C DNA achiments	1,012	115,819	7.11	(47)
h) Hunun herpes virus type 6 DNA polymerase	985	113,468	7.98	
i) Channel Carfish virus DNA polymerase	913	104,955	6.66	(48)
j) Chlorella virus DNA polymerase	988	116,658	8.11	(49)
k) Fowlpox virus DNA polymerase	937	108,564	7.50	(50)
1) Vaccinia virus DNA polymerase	964	114,818	7.95	(51)
m) Charlstoneura biennis DNA polymerase	7U*			
n) Autographa californica nuclear polyhedrosis	984	114,337	8.35	(52)
virum (AcMNPV) DNA polymerase	704			
o) Lymantria dispar nuclear polyhedrosis virus	1,013	115,921	9.08	(53)
	1 (113	113,741		

p) Adenovirus-2 DNA polymerase* q) Adenovirus-7 DNA polymerase* r) Adenovirus-12 DNA polymerase* 6. Eukaryotic linear DNA plasmid encoded DNA polymerases a) S-1 maize DNA polymerase* b) kalilo neurospora intermedia DNA polymerase* c) pAI2 Ascobolus immersus DNA polymerase* d) pCLK1 Claviceps purpurea DNA polymerase* e) maranhar neurospora crassa DNA polymerase* f) pEM Agaricus bitorquis DNA polymerase*	1,056 1,122 1,053 917 970 1,202 1,097 1,021 †797 995	120,431 128,648 120,863 105,935 112,902 138,279 126,627 119,074 †91,922 116,345	6.65 6.73 6.86 8.62 9.71 10.10 8.76 9.62 †8.24 8.04	(54) (55) (56) (57) (58) (59) (60) (61) (62) (63)
g) pGKL1 Kluyveromyces lactis DNA polymerase* h) pGKL2 Kluyveromyces lactis DNA polymerase* i) pSKL Saccharomyces kluyveri DNA polymerase* C. Family C DNA polymerases	993 994 999	117,560 117,544	8.33 9.79	(64) (65)
Bacterial replicative DNA polymerases	4460	129,903	5.04	(66)
a) E. coli DNA polymerase III α subunit	1160	130,118	5.05	(67)
b) S. typhimurium DNA polymerase III α subunit	1160		5.23	(68)
c) Bacillus subtilis DNA polymerase III	1437	162,648	2.22	\>
2. E.coli dnaQ (MutD) E.coli DNA polymerase III \(\epsilon \) subunit	243	27,099	5.68	(69,70)

- 4. Lopez, P., Martinez, S., Diaz, A., Espinosa, M. and Lacks, S. A. (1989) J. Biol. Chem. 264, 4255-4263.
- 5. Lawyer, F. C., Stoffel, S., Saiki, R. K., Myambo, K., Drummond, R. and Gelfand, D. H. (1989) J. Biol. Chem. 264, 6427-6437.
- Akhmetzjanov, A.A. and Vakhitov, V.A. (1992) Nucl. Acids Res. 20, 5839. 7. Leavitt, M. C. and Ito, J. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 4465-4469.
- 8. Dunn, J. J. and Studier, F. W. (1983) J. Mol. Biol. 166, 477-535.
- 9. Scarlato, V. and Gargano, S. (1992) Gene 118, 109-113.
- 10. Rådén, B. and Rutberg, L. (1984) J. Virol. 52, 9-15.
- 11. Foury, F. (1989) J. Biol. Chem. 264, 20552-20560.
- 12. Ito, J. and Braithwaite, D.K. (1990) Nucl. Acids Res. 18, 6716.
- 13. Blanco, L., Bernad, A. and Salas, M. (1991) Nucl. Acids Res. 19, 955.
- 14. Hahn, S. and Rüger, W. (1989) Nucl. Acids Res. 17, 6729. 15. Hollingsworth, H.C. and Nossal, N.G. (1991) J. Biol. Chem. 266,
- 1888-1897. 16. Kaliman, A.V., Krutilina, A.I., Kryukov, V.M. and Bayev, A.A. (1986) FEBS Lett. 195, 61-64.
- 17. Iwasaki, H., Ishino, Y., Toh, H., Nakata, A. and Shinagawa, H. (1991) Mol. Gen. Genet. 226, 24-33.
- 18. Jung, G., Leavitt, M. C., Hsieh, J.-C. and Ito, J. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 8287-8291.
- 19. Savilahti, H. and Bamford D.H. (1987) Gene 57, 121-130.
- 20. Yoshikawa, H. and Ito, J. (1982) Gene 17, 323-335.
- 21. Matsumoto, K., Takano, H., Kim, C. I. and Hirokawa, H. (1989) Gene 84, 247-255.
- 22. Spicer, E. K., Rush, J., Fung, C., Reha-Krantz, L. J., Karam, T. D. and Konigsberg, W. H. (1988) J. Biol. Chem. 263, 7478-7486.
- 23. Perler, F.B., Comb, D.G., Jack, W.E., Moran, L.S., Qiang, B., Kucera, R.B., Benner, J., Slatko, B.E., Nwankwo, D.O., Hempstead, S.K., Carlow, C.K.S. and Jannasch, H. (1992) Proc. Natl. Acad. Sci. USA 89, 5577 – 5581.
- 24. Cline, J.M., Scott, B.R., Adams, M.W.W., Sorge, J.A. and Mathur, E.J. Gene (in press).
- 25. Pisani, F.M., De Martino, C. and Rossi, M. (1992) Nucl. Acids Res. 20,
- 2711 2716. 26. Wong, S. W., Wahl, A. F., Yuan, P.-M., Arai, N., Pearson, B. E., Arai, K.-i., Korn, D., Hunkapiller, M. W. and Wang, T.S.-F. (1988) EMBO J. 7, 37-47.
- 27. Pizzagalli, A., Valsasnini, P., Plevani, P. and Lucchini, G. (1988)Proc. Natl.
- Acad. Sci. U.S.A. 85, 3772-3776. 28. Damagnez, V., Tillit, J., de Recondo, A.-M. and Baldacci, G. (1991) Mol.
- Gen. Genet. 226, 182-189. 29. Hirose, F., Yamaguchi, M., Nishida, Y., Masutani, M., Miyazawa, H., Hanaoka, F. and Matsukage, A. (1991) Nucl. Acids Res. 19, 4991-4998.
- 30. Leegwater, P.A.J., Strating, M., Murphy, N.B., Kooy, R.F., van der Vliet, P.C. and Overdulve, J.P. (1991) Nucl. Acids Res. 19, 6441-6447.
- 31. Chung, D.W., Zhang, J., Tan C.-K., Davie, E.W., So, A.G. and Downey, K.M. (1991) Proc. Natl. Acad. Sci. USA 88, 11197-11201.

- 32. Yang, C.-L., Chang, L., S., Zhang, P., Hao, H., Zhu, L., Tommey, N.L. and Lee, M.Y.W.T. (1992) Nucl. Acids Res. 20, 735-745.
- 33. Zhang, J., Chung, D.W., Tan, C.-K., Downey, K.M., Davie, E.W. and So, A.G. (1991) Biochemistry 30, 11742-11750.
- 34. Morrison, A. and Sugino, A. (1992) Nucl. Acids Res. 20, 375.
- 35. Pignéde, G., Bouvier, D., de Recondo, A.-M. and Baldacci, G. (1991) J. Mol. Biol. 222, 209-218.
- 36. Ridley, R.G., White, J.H., McAleese, S.M., Goman, M., Alano, P., de Vries, E. and Kilbey, B.J. (1991) Nucl. Acids Res. 19, 6731-6736.
- Morrison, A., Araki, H., Clark, A.B., Hamatake, R.K. and Sugino, A. (1990) Cell 62, 1143-1151.
- 38. Morrison, A., Christensen, R. B., Alley, J., Beck, A. K., Bernstine, E. G., Lemontt, J. F. and Lawrence, C. W. (1989) J. Bacteriol. 171, 5659-5667.
- 39. Gibbs, J. S., Chiou, H. C., Hall, J. D., Mount, D. W., Retondo, M. J., Weller, S. K. and Coen, D. M. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 7969-7973.
- Telford, E.A., Watson, M.S., McBride, K. and Davison, A.J. (1992) Virology 189, 304-316.
- 41. Davison, A. J. and Scott, J. E. (1986) J. Gen. Virol. 67, 1759-1816.
- 42. Baer, R., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Séguin, C., Tuffnell, P.S. and Barrell, B.G. (1984) Nature 310, 207-211.
- 43. Albrecht, J.-C. and Fleckenstein, B. (1990) Virology 174, 533-542.
- 44. Kouzarides, T., Bankier, A. T., Satchwell, S. C., Weston, K., Tomlinson, P. and Barrell, B. G. (1987) J. Virol. 61, 125-133.
- 45. Elliott, R., Clark, C. Jaquish, D. and Spector, D.H. (1991) Virology 185, 169 - 186.
- Teo, I.A., Griffin, B.E. and Jones, M.D. (1991) J. Virol. 65, 4670-4680.
- 47. Davison, A.J. (1992) Virology 186, 9-14.
- 48. Grabherr, R., Strasser, P. and Van Etten, J.L. (1992) Virology 188, 721-731.
- 49. Binns, M. M., Stenzler, L., Tomley, F. M., Campbell, J. and Boursnell, M. E. G. (1987) Nucl. Acids Res. 15, 6563-6573.
- 50. Eari, P. L., Jones, E. V. and Moss, B. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 3659-3663.
- 51. Mustafa, A. and Yuen, L. (1991) DNA Seq. 2, 39-45.
- 52. Tomalski, M. D., Wu, J. and Miller, L. K. (1988) Virology 167, 591-600.
- Bjornson, R.M. and Rohrmann, G.F. (1992) J. Gen. Virol. 73, 1499-1504.
- Gingeras, T. R., Sciaky, D., Gelinas, R. E., Bing-Dong, J., Yen, C. E., Kelly, M. M., Bullock, P. A., Parsons, B. L., O'Neill, K. E. and Roberts, R. J. (1982) J. Biol. Chem. 257, 13475-13491.
- 55. Engler, J.A., Hoppe, M.S. and van Bree, M.P. (1983) Gene 21, 145-159.
- 56. Shu, L., Hong, J.S., Wei, Y.-f. and Engler, J.A. (1986) Gene 46, 187-195. 57. Paillard, M., Sederoff, R. R. and Levings, C. S. III (1985) EMBO J. 4,
- 1125-1128.
- 58. Chan, B.S.-S., Court, D.A., Vierula, P.J. and Bertrand, H. (1991) Curr. Genet. 20, 225-237.

- 59. Kempken, F., Meinhardt, F. and Esser, K. (1989) Mol. Gen. Genet. 218,
- 60. Oeser, B. and Tudzynski, P. (1989) Mol. Gen. Genet. 217, 132-140.
- 61. Court D.A. and Bertrand, H. (1992) Curr. Genet. 22, 385-397.
- 62. Robison, M.M., Royer, J.C. and Horgen, P.A. (1991) Curr. Genet. 19,
- 63. Stark, M.J.R., Mileham, A.J., Romanos, M.A. and Boyd, A. (1984) Nucl. Acids Res. 12, 6011-6030.
- 64. Tommasino, M., Ricci, S. and Galeotti, C. L. (1988) Nucl. Acids Res. 16,
- 65. Hishinuma, F. and Hirai, K. (1991) J. Gen. Genet. 226, 97-106.
- 66. Tomasiewicz, H. G. and McHenry, C. S. (1987) J. Bacteriol. 169,
- 67. Lancy, E. D., Lifsics, M. R., Munson, P. and Maurer, R. (1989) J. Bacteriol.
- Hammond, R. A., Barnes, M. H., Mack, S. L., Mitchener, J. A. and Brown, N. C. (1991) Gene 98, 29-36.
 Maki, H., Horiuchi, T. and Sekiguchi, M. (1983) Proc. Natl. Acad. Sci. USA 80, 7137-7141.
- 70. Joyce, C.M., Ollis, D.L., Rush, J., Steitz, T.A., Konigsberg, W.H. and Grindley, N.D.F. (1986) in 'Protein structure, folding and design (UCLA Symposia on Molecular and Cellular Biology, Vol 32)', D. Oxender, Ed., pp 197-205, Alan R. Liss.
- 71. Felsenstein, J. (1989) Cladistics 5, 164-166.

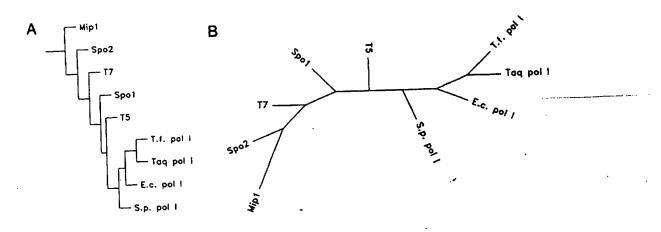


Figure 2A. Phylogenetic phenogram tree produced from the alignment of Family A DNA polymerases in Figure 1A using only the following conserved sequence blocks of the DNA polymerase domain: 798-814, 877-998, 1047-1090, 1104-1123, 1131-1158, 1175-1206, 1236-1251, 1284-1305, 1322-1340, and blocks of the DNA polymerase domain: 798-814, 877-998, 1047-1090, 1104-1123, 1131-1158, 1175-1206, 1236-1251, 1284-1305, 1322-1340, and 1365-1379. B. Unrooted phylogenetic tree of the Family A DNA polymerases produced as in Figure 2A plotted by a different method.

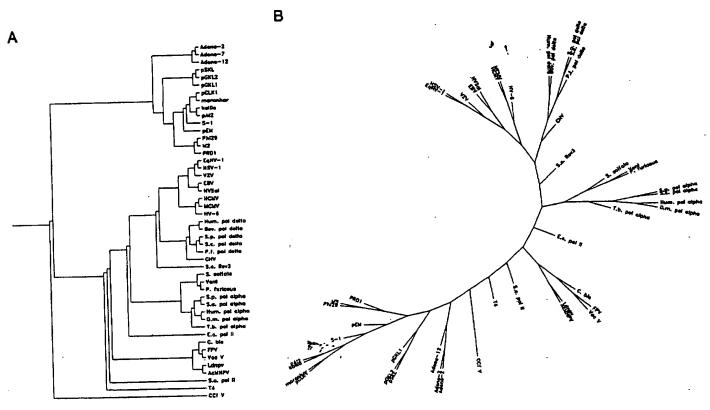


Figure 3A. Phylogenetic phenogram tree produced from the alignment of Family B DNA polymerases in Figure 1B using only the following conserved sequence blocks of the DNA polymerase domain: 1407-1760, 1885-1901, 1956-1990, 2081-2100, 2181-2210, and 2280-2320. B. Unrooted phylogenetic tree of the Family B DNA polymerases produced as in Figure 3A plotted by a different method.

ACKNOWLEDGEMENTS

The authors are grateful to all the individuals who have kindly communicated their unpublished sequences to be included in this compilation. This work was supported in part by Grant GM28013 from the NIH to J.I.

REFERENCES

- 1. Ito, J. and Braithwaits, D.K. (1991) Nucl. Acids Res. 19, 4045-4057.
- 2. Wheelis, M.L., Kandler, O. and Woese, C.R. (1992) Proc. Natl. Acad. Sci. USA 89, 2930-2934.
- 3. Joyce, C. M., Kelley, W. S. and Grindley N. D. F. (1982) J. Biol. Chem. 257, 1958-1964.

Compilation and alignment of DNA polymerase sequences

Junetsu Ito and Dan K.Braithwaite

Department of Microbiology and Immunology, College of Medicine, University of Arizona Health Sciences Center, Tucson, AZ 85724, USA

Received May 29, 1991; Revised and Accepted July 9, 1991

INTRODUCTION

More than 40 different DNA polymerases, including some putative DNA polymerase sequences deduced from nucleotide sequence data, have recently been reported (1-39). The amino acid sequences of these DNA polymerases have been aligned and partial homologous regions identified by many investigators (2-4,9,10,12-25,27-36,42-51). Based on the segmental amino acid sequence similarities, DNA polymerases have been classified into two major groups; E. coli DNA polymerase I-Type and eukaryotic DNA polymerase α -Type (14,44,47,48,51), or family A DNA polymerases and family B DNA polymerases (4,9,50). As the number of DNA polymerase sequences increases, the classification of DNA polymerases becomes increasingly ambiguous. For example, DNA polymerase delta of yeast was shown to have amino acid sequence similarity to the α -Type DNA polymerases (17). It has become necessary to establish a unified classification of DNA polymerases. Here we pr pose to classify DNA polymerases into families A, B, and C (Figure 1: A, B, and C), according to the amino acid sequence homologies with E. coli DNA polymerases I, II, and III, respectively. As new and different prokaryotic and eukaryotic DNA polymerases are identified, the number of families can easily be expanded by using additional letters of the alphabet (i.e., D, E, etc.).

The bacterium *E. coli* (strain K12) contains three distinct DNA polymerases I, II, and III (52). *E. coli* DNA polymerase I, the first DNA polymerase discovered, is specified by the *pol*A gene (52). *E. coli* DNA polymerase II, encoded by the *pol*A gene, was recently sequenced and found to be identical to the *din*A gene, a DNA damage inducible gene whose expression is regulated by the SOS system in *E. coli* (8,53). Amino acid sequence alignment shows that *E. coli* DNA polymerase II has significant homology with family B (α -Type) DNA polymerases (8,53,54).

E. coli DNA polymerase III is a multisubunit enzyme encoded by various dna genes (55); the DNA polymerizing α -subunit encoded by the polC (dnaE) gene (56) and the $3' \rightarrow 5'$ exonuclease performing ϵ -subunit encoded by the dnaQ gene (57). The α -subunit of E. coli DNA polymerase III exhibits an extensive hom logy with the corresponding α -subunit of Salmonella typhimurium DNA polymerase III (35); and both show significant homology to Bacillus subtilis DNA polymerase III, a single-polypeptide encoded by the polC gene (36).

In summary, family A DNA polymerases are named for their homology to the product of the polA gene encoding E. coli DNA polymerase I; family B DNA polymerases are named for their

homology to the product of the *polB* gene encoding *E. coli* DNA polymerase II; and family C DNA polymerases are named for their homology to the product of the *polC* gene encoding *E. coli* DNA polymerase III.

The eukaryotic DNA polymerase β , the smallest known DNA polymerase, does not have homology with those of any of the DNA polymerase families described above. Instead, DNA polymerase β has homology with terminal transferases (37). This β group we will call family X (Figure 1D). The classification and original reference(s) for the amino acid sequences of each DNA polymerase are shown in Table 1.

All of the family A DNA polymerases, except for yeast mitochondrial DNA polymerase I, are prokaryotic and are very sensitive to dideoxynucleotide inhibitors, and therefore are useful enzymes for DNA sequencing by the chain-termination method (58). The family A DNA polymerases are resistant to aphidicolin. The family B DNA polymerases are quite extensive in number and variety. Most of the family B DNA polymerases, if not all, are sensitive to aphidicolin and relatively resistant to dideoxynucleotide inhibitors. Most of the family B DNA polymerases, except for pAI2 (33) and yeast DNA polymerase II (16), contain the highly conserved amino acid sequence motif YGDTD, which has been suggested to form part of the dNTP binding site. Amino acid substitutions in this conserved sequence resulted in defects in the DNA polymerase activity without affecting the $3' \rightarrow 5'$ exonuclease activity (59,60,61). The family C DNA polymerases are major bacterial replicative DNA polymerases which do not have appreciable homology with those of family A and B DNA polymerases. B. subtilis DNA polymerase III is a single polypeptide that is highly sensitive to hydroxyphenylazouracil (62). It is anticipated that the number of sequenced family C DNA polymerases will increase rapidly, since all of the aerobic bacteria may contain a member of this family of DNA polymerases.

SEQUENCE ALIGNMENT

The 37 complete DNA polymerase sequences and 3 complete terminal deoxynucleotidyltransferase (TDT) sequences are listed in 4 groups; the family A DNA polymerases, the family B DNA polymerases, the family C DNA polymerases, and family X DNA polymerases (including TDTs). In order to limit the space needed for the alignment, we omitted DNA polymerase sequences that are very similar to the prototype DNA polymerase. The DNA polymerases not shown include: herpes virus type-2 (63),

adenovius type-5 (64), bacteriophage T3 (65), and bacteriophage PZA (66)L

ACCURACY OF SEQUENCE DATA

Whenever a sequence ambiguity existed in a published sequence, we connected the authors to obtain the updated sequence information. We found that a few published amino acid sequences differ at me or more positions from their GenBank/EMBL entry. Again, we have communicated with the primary author to confirm the correct sequences.

The multiple alignment of the amino acid sequences was obtained by a series of pairwise alignments combined and adjusted by eye in larger and larger subsets of similar sequences. The process at combining and adjusting by eye was aided by modified versions at the MOTIF program (67) and the ALIGN program (68). The GAP and BESTFIT programs, from UWGCG (University of Wisconsin Genetic Computer Group) (69), initially generates the pairwise alignments, adjusted for maximum alignmen that allowed for a considerable number of gaps. We then compressed these alignments by eye to give a more contiguous alignment. The alignment of the sequences for optimal similarity is straightforward in the areas of relatively conserved structure, but is much more arbitrary in the more varied sequence areas. The alignment of the varied areas should therefore be regarded is less than optimal in view of the difficulties concerned with multiple alignments in these areas.

Finally, we invite further correction from readers, and welcome suggester revisions and alternative alignments.

ACKNOWLEDGEMENTS

This work was supported by grant GM28013 from the National Institutes of Health and by grant NP-704 from the American Cancer Society.

REFERENCES

- 1. Joyce, C.M., Kelley, W.S. and Grindley N.D.F. (1982) J. Biol. Chem. 257, 1958-964.
- Lopez P., Martinez, S., Diaz, A., Espinosa, M. and Lacks, S.A. (1989) J. Bin. Chem. 264, 4255-4263.
- 3. Lawe F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R. and Gelfinst. D.H. (1989) J. Biol. Chem. 264, 6427-6437.
- Leavit. M.C. and Ito, J. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 4463--469.
- 5. Dunn. I.J. and Studier, F.W. (1983) J. Mol. Biol. 166, 477-535.
- 6. Radez B. and Rutberg, L. (1984) J. Virol. 52, 9-15.
- 7. Four. F. (1989) J. Biol. Chem. 264, 20552-20560.
- 8. Iwasas, H., Ishino, Y., Toh, H., Nakata, A. and Shinagawa, H. (1991) Mol. Gen. Genet. 226, 24-33.
- 9. Jung. Z. Leavitt, M.C., Hsieh, J.-C. and Ito, J. (1987) Proc. Natl. Acad. Sci. LS.4. 84, 8287-8291.
- Savison. H. and Bamford D.H. (1987) Gene 57, 121-130.
- 11. Yoshisawa, H. and Ito, J. (1982) Gene 17, 323-335.
- 12. Massimoto, K., Takano, H., Kim, C.I. and Hirokawa, H. (1989) Gene 84,
- 13. Spice: E.K., Rush, J., Fung, C., Reha-Krantz, L.J., Karam, J.D. and Konissierg, W.H. (1988) J. Biol. Chem. 263, 7478-7486.
- Work S.W., Wahl, A.F., Yuan, P.-M., Arai, N., Pearson, B.E., Arai, K.-i. Korn, D., Hunkapiller, M.W. and Wang, T.S.-F. (1988) EMBO J.
- 15. Pizzani, A., Valsasnini, P., Plevani, P. and Lucchini, G. (1988) Proc. Nati scad. Sci. U.S.A. 85, 3772-3776.
- 16. Morraga. A., Araki, H., Clark, A.B., Hamatake, R.K. and Sugino, A. (1990) Cell & 1143-1151.

- 17. Boulet, A., Simon, M., Faye, G., Bauer, G.A. and Burgers, P.M.J. (1989) EMBO J. 8, 1849-1854.
- 18. Morrison, A., Christensen, R.B., Alley, J., Beck, A.K., Bernstine, E.G., Lemontt, J.F. and Lawrence, C.W. (1989) J. Bacteriol. 171, 5659-5667.
- 19. Gibbs, J.S., Chiou, H.C., Hall, J.D., Mount, D.W., Retondo, M.J., Weller,
- S.K. and Coen, D.M. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 7969-7973. 20. Kouzarides, T., Bankier, A.T., Satchwell, S.C., Weston, K., Tomlinson,
- P. and Barrell, B.G. (1987) J. Virol. 61, 125-133.
- 21. Baer, R., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Séguin, C., Tuffnell, P.S. and Barrell, B.G. (1984) Nature 310, 207-211.
- 22. Davison, A.J. and Scott, J.E. (1986) J. Gen. Virol. 67, 1759-1816.
- 23. Binns, M.M., Stenzler, L., Tomley, F.M., Campbell, J. and Boursnell, M.E.G. (1987) Nucl. Acids Res. 15, 6563-6573.
- 24. Earl, P.L., Jones, E.V. and Moss, B. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 3659-3663.
- 25. Tomalski, M.D., Wu, J. and Miller, L.K. (1988) Virology 167, 591-600.
- 26. Gingeras, T.R., Sciaky, D., Gelinas, R.E., Bing-Dong, J., Yen, C.E., Kelly, M.M., Bullock, P.A., Parsons, B.L., O'Neill, K.E. and Roberts, R.J. (1982) J. Biol. Chem. 257, 13475-13491.
- 27. Engler, J.A., Hoppe, M.S. and van Bree, M.P. (1983) Gene 21, 145-159.
- 28. Shu, L., Hong, J.S., Wei, Y.-f. and Engler, J.A. (1986) Gene 46, 187-195.
- 29. Paillard, M., Sederoff, R.R. and Levings, C.S. III (1985) EMBO J. 4, 1125 - 1128.
- 30. Stark, M.J.R., Mileham, A.J., Romanos, M.A. and Boyd, A. (1984) Nucl. Acids Res. 12, 6011-6030.
- Tommasino, M., Ricci, S. and Galeotti, C.L. (1988) Nucleic Acids Res. 16, 5863-5878.
- 32. Oeser, B. and Tudzynski, P. (1989) Mol. Gen. Genet. 217, 132-140.
- Kempken, F., Meinhardt, F. and Esser, K. (1989) Mol. Gen. Genet. 218, 523 - 530.
- Tomasiewicz, H.G. and McHenry, C.S. (1987) J. Bacteriol. 169, 5735-5744.
- Lancy, E.D., Lifsics, M.R., Munson, P. and Maurer, R. (1989) J. Bacteriol. 171, 5581-5586.
- Hammond, R.A., Barnes, M.H., Mack, S.L., Mitchener, J.A. and Brown, N.C. (1991) Gene 98, 29-36.
- Matsukage, A., Nishikawa, K., Ooi, T., Seto, Y. and Yamaguchi, M. (1987) J. Biol. Chem. 262, 8960-8962.
- 38. Abbotts, J., SenGupta, D.N., Zmudzka, B., Widen, S.G., Notario, V. and Wilson, S.H. (1988) Biochemistry 27, 901-909.
- 39. SenGupta, D.N., Zmudzka, B.Z., Kumar, P., Cobianchi, F., Skowronski, J. and Wilson, S.H. (1986) Biochem. Biophys. Res. Comm. 136, 341-347.
- Peterson, R.C., Cheung, L.C., Mattaliano, R.J., White, S.T., Chang, L.M.S. and Bollum F.J. (1985) J. Biol. Chem. 260, 10495-19502.
- Koiwai, O., Yokota, T., Kageyama, T., Hirose, T., Yoshida, S. and Arai, K.-i. (1986) Nucl. Acids Res. 14, 5777-5792.
- 42. Argos, P., Tucker, A.D. and Phillipson, L. (1986) Virology 149, 208-216.
- 43. Larder, B.A., Kemp, S.D. and Darby, G. (1987) EMBO J. 6, 160-175. 44. Hall, J.D. (1988) Trends Genet. 4, 42-46.
- 45. Reha-Krantz, L.J. (1988) J. Mol. Biol. 202, 711-724.
- 46. Bernad, A., Zaballos, A., Salas, M. and Blanco, L. (1987) EMBO J. 6, 4219-4225
- 47. Wang, T.S.-F., Wong, S.W. and Korn, D. (1989) FASEB 3, 14-21.
- Bernad, A., Blanco, L., Lazaro, J.M., Martin, G. and Salas, M. (1989) Cell 59, 219-228.
- Polesky, A.H., Steitz, T.A., Grindley, N.D.F. and Joyce, C.M. (1990) J. Biol. Chem. 24, 14579-14591.
- 50. Ito, J. and Braithwaite, D.K. (1990) Nucl. Acids Res. 18, 6716.
- 51. Blanco, L., Bernad, A. and Salas, M. (1991) Nucl. Acids Res. 19, 955.
- 52. Kornberg, A. (1974) DNA replication. W.H. Freeman and Co., San Francisco.
- 53. Bonner, C.A., Hays, S., McEntee, K. and Goodman M.F. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 7663-7667.
- 54. Chen, H., Lawrence, C.B., Bryan, S.K. and Moses, R.E. (1990) Nucl. Acids Res. 18, 7185-7186.
- 55. McHenry, C.S.(1985) Mol. Cell. Biochem. 66, 71-85.
- 56. Shepard, D., Oberfelder, R.W., Welch, M.W. and McHenry, C.S. (1984) J. Bacteriol. 158, 455-459.
- 57. Scheuermann, R., Tam, S., Burgers, P.M.J., Lu, C. and Echols, H. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 7085-7089. 58. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci.
- U.S.A. 74, 5463-5467.
- Dorsky, D.I. and Crumpacker, C.S. (1990) J. Virol. 64, 1394-1397. 60. Bernad, A., Lázaro, J.M., Salas, M. and Blanco, L. (1990) Proc. Natl. Acad.
- Sci. U.S.A. 87, 4610-4614.

- 61. Jung, G., Leavitt, M.C., Schultz, M. and Ito, J. (1990) Biochem. Biophys. Res. Comm. 170, 1294-1300.
- 62. Neville, M.N. and Brown, N.C. (1972) Nature New Biol. 240, 80-82.
- 63. Tsurumi, T., Maeno, K. and Nishiyama, Y. (1987) Gene 52, 129-137.
- 64. Dekker, B.M.M. and Van Ormondt, H. (1984) Gene 27, 115-120.
- 65. Beck, P.J., Gonzalez, S., Ward, C.L. and Molineux, I.J. (1989) J. Mol. Biol. 210, 687-701.
- 66. Paces, V., Vlcek, C., Urbanek, P. and Hostomsky, Z. (1985) Gene 38, 45-56.
- 67. Smith, H.O., Annau, T.M. and Chandrasegaran, S. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 826-830.
- 68. Doolittle, R.F. and Feng, D.-F. (1990) In Doolittle, R.F. (ed.), Methods in Enzymol. - Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences. Academic Press, New York, Vol. 183, pp. 659-669.
- 69. Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucl. Acids Res. 12, 387 - 395.

Cl

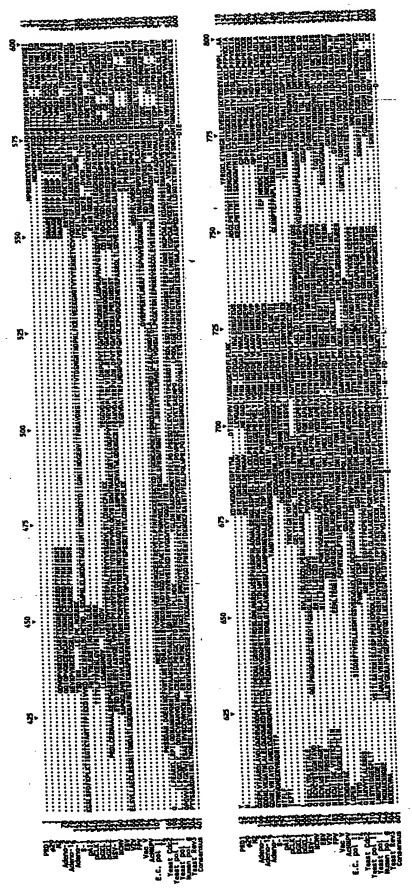
Fami	ly A D	NA polymerases	
1.	Bacte	erial DNA polymerases	eferences
	a)	E. coli DNA polymerase I	(1)
	b)	Strentococcus pneumoniae DNA polymerase I	(2)
	c)	Thermus aquaticus DNA polymerase I	(3)
2.	Bact	eriophage DNA polymerases	
	a)	TS DNA polymerase	(4)
	b)	T7 DNA polymerase	(5) (6)
	c)	Spo2 DNA polymerase	(0)
3.	Mite	schondrial DNA polymerase	
		Yeart mitochondrial DNA polymerase (MIP1)	(7)
Fam	ily B I	DNA polymerases	
1.	Baci	erial DNA polymerases	
		E. coli DNA polymerase II	(8)
2.	Baci	teriophage DNA polymerases	
	a)	PRD1 DNA polymerase*	(9,10) (11)
	Þ)	629 DNA polymerase*	(11)
	c) d)	M2 DNA polymerase T4 DNA polymerase	(13)
	•	·	
3.	Enk	aryotic DNA polymerases	(14)
	a)	Human DNA polymerase alpha	(15)
	p)	Yeast DNA polymerase I Yeast DNA polymerase II	(16)
	c) d)	Yeast DNA polymerase III (delta)	(17)
	e)	Yeast DNA polymerase Rev3	(18)
4.	Vlr	al DNA polymerases	
	a)	Herpes-1 DNA polymerase	(19)
	b)	Human cytomegalovirus DNA polymerase	(20) (21)
	c)	Epstein-Barr virus DNA polymerase Varicella-Zoster virus DNA polymerase	(22)
	d) e)	Fowlpox virus DNA polymerase	(23)
	ŋ	Vaccinia virus DNA polymerase	(24)
	g)	Autographa californica nuclear	. /263
		polyhedrosis virus (AcMNPV) DNA polymerass Ademovirus-2 DNA polymerase*	25) (26)
	P)	Adenovirus-7 DNA polymerase*	(27)
	Ŋ	Adenovirus-12 DNA polymerase*	(28)
5.	Eol	caryotic linear DNA plasmid encoded DNA polymerases	
	a)	S-1 maize mitochondrial DNA polymerase*	(29)
	p).	Kluyveromyces lactis plasmid pGKL1 DNA polymerase Kluyveromyces lactis plasmid pGKL2 DNA polymerase	* (30) * (31)
	c) d)	Claviceps purpures plasmid pCLK1 DNA polymerase*	(32)
	e)	Ascobolus immersus plasmid pAI2 DNA polymerase*	(33)
Faz	ыйу С	DNA polymerases	
	Bac	cterial replicative DNA polymerases	
	a)	E. coii DNA polymerase III a subunit Salmonella typhimurium DNA polymerase III a subun	(34) it (35)
	b) c)	Bacillus subdilis DNA polymerase III	(36)
Fac	mily X	DNA polymerases	
	a)	Rat DNA polymerase β	(37)
	b)	Human DNA polymerase β	(38,:
	c)	Human terminal decaynucleotidyltransferase (TdT) Bovine terminal decaynucleotidyltransferase (TdT)	(40) (41)
	d)	Mouse terminal decrynacieotidyltransferase (TdT)	(41)

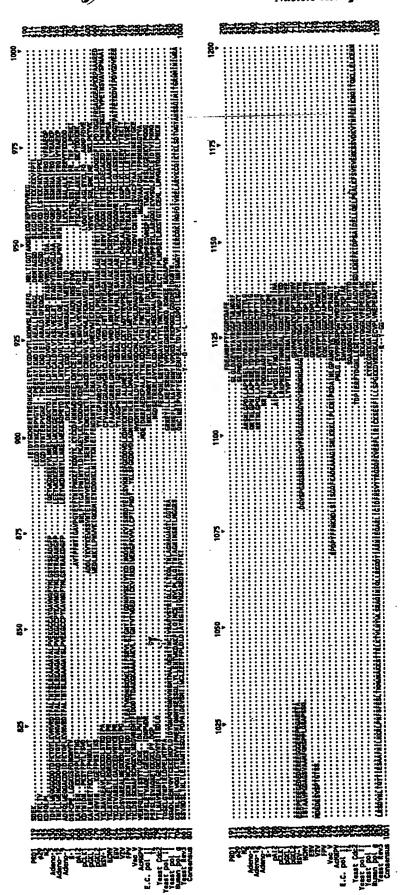
Table 1. The main families and subclassifications of DNA polymerases. Those DNA polymerases marked with a star (*) are protein-primed DNA polymerases.

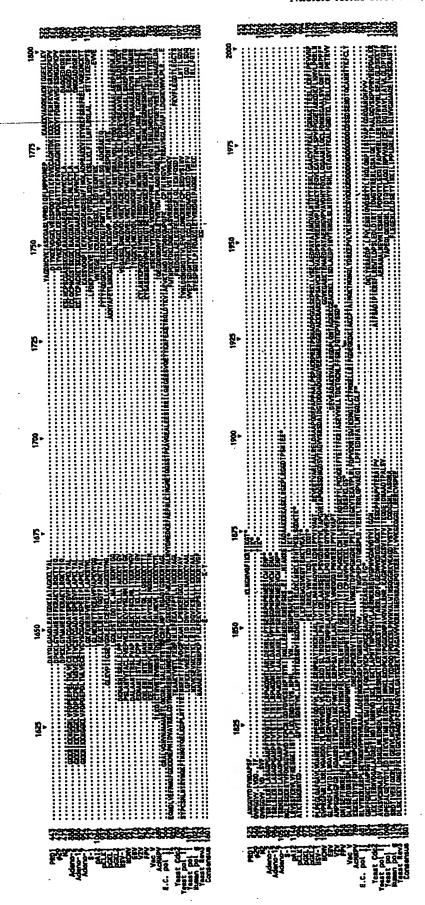
MAC TO THE		3 22572718 3 25572718		25335828	**************************************	25 25 25 25 25 25 25 25 25 25 25 25 25 2	***
22-8-80 0-8-80 0-8-80 0-8-80 0-8-80 0-8-80	400 AVIDVAN VVIDESC VV	Froiries 600 1 - IASL'60 11 - IASL'60 12 - IASL'60 13 - IASL'60 14 - IASL'60 15 - IASL'60 16 - IASL'60 17 - IASL'60 18	800 Krolotilfer Krolotilfer Krolotilfer Krolotilfer Adiktilfer Adiktilfer	1000 ILAF IRAF PAVE FRARE	1200 SPIGGTAND SPIGGTAND SPIGGSAND ALIUGSAND ILIUMATAND ILIUMATAND ILIUMATAND ILIUMATAND ILIUMATAND ILIUMATAND ILIUMATAND ILIUMATAND ILIUMATAND	15. 15. 15. 15. 15. 15. 15. 15. 15. 15.	
1000 1000 1000 1000	SFALEPGYA TOAL BYSSAA SI SHORTOGY VI YOTSTEA	TFNG LREGITAP KREGITAP KREGITAP ETVOSEL	FALSSTOL FALVESP	1000	AAAERAIN Eraerain Eraerain Eraerain Eraerain Tygeriven Tygeriven	1375 Oversockovneleběli rěci skérkkočatký	3
212 222 222 222 222 222	DESCRIPTION OF THE PROPERTY OF	AAA CAAA CAAA CAAA CAAA CAAA CAAA CAAA	A Present	975 SET SOLETIVE SET SOLETIV		1375	S S S S S S S S S S S S S S S S S S S
252 	375 Govern Caftolina Afdren Caftolina Afdren Caftolina Afdren Caftolina Afdren Caftolina Aftren Caftolina	STATE OF THE STATE	ARI EKKANE LAGE EKI TOLITELAGE WALE ELYANDA TOLITE	PETER SEPTION OF THE PETER SEP	EALTTEGG	92005	KAVISC
rve III	CLECLYTS GLEDLYTS THAY (GPV	AVPORELIN CYTOMODYO LGYRLGEPKO	LACACLERY LYCKLICA LYCKLICA PEYKOLEO PE	24444 44444 44444 44444	SSMGARR SSMFRIR SEDRGYG SEDRGYG SELDERFG		
DENCH TO THE TOPH TOPH TOPH TOPH TOPH TOPH TOPH TOPH	350 EETLKAULA 16TKAPJA 16ZEEŤÍČOV	SSO SSO LLAPAR	750 TOWNER TIRU CKETTER SERVENT CKETTER SERVENT WILKARERY TYPEN MILKARERY TYPEN MILKARERY TOWNER MILKARERY TOWNER MILKARERY TOWNER TOWNER LEPSTOTTER TIKE	250 PATENCY PA	1150 XRRELPOIN ZRRELPOIN ZRRYPPOLE ZRRYPPO Z	1350 Averoldavkriled	155d
101 101 101 101 101 101 101 101 101 101	TOWYTHE AFLSKTLAT	COXNOLOGY CALROLOGY TOUREN TYLOGERECTI TYLOGERECTI TASTOR	2000 E	35		PATE OF STREET	VE AD
AREAECAGE PKLAGOOGE AKAGOOGE	IPEVTATVIS CERLINOKEI	SPICE PRICE SECRETARY CHINGS PRIVATE PRIVATE	2001/2007/4/07/1/19:16:16:16:16:16:16:16:16:16:16:16:16:16:	Falconia I I Pries I P	#6.11.12.	1350 KROJO SI KEIGODETTÄYVROVLLÖLE PETEKKAANDERIPONSKYRODER I ANCHOLDANKRILLOA BEENVAGETTESKYA TÄYVROVLLOAE PETEKKLAANDERINDOSKYTODER I ANCHOLDANKRILLOA	901,2901 611
16 Abody Abody By Abody By By B	325 325 30ETSVADE FVENTEKN	EPTSPCOCKTERYA 525 TARTERYOPETH FLETERYOPETH FLANKOOMANA TARTERYOPETH TARTERYOPE	725 725 11. YOMEGE 11. YOMEGE 11. YOMEGE 12. YOUR TEEP 13. YOUR TEEP	26 26	1125 CERCPVKLOH	NYSOLOUS CALLACTOLI THE CALLACTOR CONTROL OF THE CALLACTOR CONTROL OF T	S
L PPLLANT	akgakpaaki Penepveda	ACRIONOSL REPRINDICAL REPRINDICAL PRENCIOTAL	EKT SENGOL HEGERALL LING GERALL TOPK GET EN		PACE SECOND CONTRACTOR	VILGAEPEY ALBINGLED	CAME
PLEASE LESS AND LESS	DVEAGINA O	GENTELTOP ISPRIPACIOLISTO BENTALLOS SON BENT	IVETEPILL ERLFAHLUG KLCSLYTOV KLTETFGSV			1300 Siaciosofythywaylldaef	
POOLRAGIE	300 XYEFKRUTA HĞYANILAS	SOCINEL SECTE CONCLUCENT LTDP ISPLEP COLBS DEF 475 SOUTH STATE STATE SOUTH STATE ST	TRAGENGPPFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	ST. TOGSENO! ESCUAL I CO IVILIZACIO I CONTRACTO POR ST. P. CONTRACTOR	ML-MIP	1300 0001 00013915916	1500 THE STATE OF THE STATE OF
NAKSKAPP TKAGRAKT	AEELLG.FK	GKNTLITOP	CKKEAEVES Aleelyvel	CONT. SSTOP CONT.	CALEGE PE	1 1 1 2 1 12 1	
2011 2011 2012 2012 2014 2014 2014 2014	S EQLEVOOPA FFACKRGIG	SKKSLKONGLV 475 11EFGSLLNEFF	ACREMENTAN	875 875 875 875 875 877 877 877 877 877	ALTINESSTANSA 1075 107	1275 1275 1275 1275 1275 1275 1275 1275	475 100Enicici
AAVVFDAKE NI LVAFDAKE	27 CTDVELELTO VICYONEPFI	GOOLINELSKELONGLUGKNI 475 GEBRICEREFESTINFFGLES	IIASFGLCSR 6 5 5AMLLAKQE	MATGERIA MATGERIA MATGERIA MATGERIA MATGERIA MATGERIA MATGERIA	LANGALOIN 1-6-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1	1 PECPTODA P	
ALLENGENERS	r. syaLATI	SCGGKDEVE SKRREPORE	FLOTOS LI	CAL COST	SAFELSE SOFE SOFELSE SOFE SOFE SOFE SOFE SOFE SOFE SOFE SO	MENCIALO MERCIALO MERCIALO MERCIALO MERCIALO MEN	SAIPEKGIL
TAGENTALES	250 KLEGMKEVA HITIGFÉW	450 450 450 801 801 801 801 801 801 801 801 801 80	650 650 WGYTT FWSE	650 650 650 650 651 651 1051 1051 1051 1	1050 1050 NWFGNYOR NWFGNYOR NFFERNYO	1250 1250 1250 1250 1250 1250 1250 1250	1430 LVANGOTT TERETAND
SAGE PTOA	250 Frcaktmaaki compevat SLPVCSKKH I 161 Pelid	MPVGTGYLG TDRCTLANY TDRLVGFA TDRLVGFA TOVPALTKL	YORARVIE YOUNG	SKCILPTKL LEYRGLAKI LOYRCIAKI LOYRCIAKI LOYRCIAKI LEFRKITKL	TERCINCE SECRESAL/ FAMERAKE FPALROAKE ELPTROMAKE ELPTROMAKE GEPTROMAKE	WEVERDOW WAYNED IN WOOTH FEEL	,000 000 100 100 100
944 944 1264 1264	EPER STATE OF STATE O	KNAAEAPEI KVCONL.K ELGENY.B HOOLKL.K	TANKE TANKE	VSFACLKSI. LOTPLPKVI FAID VEKTI TOND JAKTI.	TINGE TO THE TOTAL	A STATE OF THE STA	M-VHOE!
25 FYL YRATHA. SVAFRAFFAL KLAYRTFHAI	225 LGGLDTLYA HGSLEGIYEI	KPLLEDEKA KPLLEDEKA KPLLEDEKA KPLLEDEKA GKILDOSEHIA GKILDOSEHIA GKILDOSEHIA GKILDOSEHIA GKILDOSEHIA	ALVALISTA 625 625 625 625 625 625 625 625	825 825 SEEVLEELA SOOVLEALE DAEALWELA DOEVLEOVE	CLOCTC. V. CLOCTC. V. 1025 ANEVFG. PLE ASIMFFYPE ASIMFYP	99 V 1225 10 V 10 V	25 25 25 25 25 25 25 25 25 25 25 25 25 2
######################################	KTABALLOG KTGIKLLE KTARKLEE	VARRAPIENT SONLSEES	AGGOTLPHE COLMICHED COVERTING COVERT	RAIKHIS SOVEDKIFF CTPGGAPST CTPGGAPST TEKTGY ST TEKTGY S	ECCD (MRAT) ECCD (MRAT) ECCD (MRAT) ECCD (MRET) KYPO (MRET) KYPO (MRET) KYPO (MRET) KYPO (MRET)	A PORT OF THE PROPERTY OF THE	14.50 14.50 14.50 14.50 14.50 14.50 14.50 14.50 14.50 14.50 14.50 14.50 14.50 14.50
TS INTERPOLATION OF THE PROPERTY OF THE PROPER	25 35 35 35 35 35 35 35 35 35 35 35 35 35	LWAIFEN ENVISED FRONTLANISTECKALENT INVGIOTICES GAROVESS LIV DOI SCORT SEES 425 426 FILV DOI SCORT SEES 426 427 427 428 429 429 420 420 420 420 420 420	000 650 650 650 650 650 650 650 650 650	A STATE LEGISTATION PROTESTAND AND AND AND AND AND AND AND AND AND	COL 1155/1961 1960 1960 1960 1960 1960 1960 1960	1225 1226 1227 1227 1228 1229 1229 1220 1220 1220 1221 1221 1222 1220 1220 1220 1221 1222 1223 1220 1220 1220 1220 1220 1221 1222 1223 1220 1220 1223 1223 1220	Nimodamajoniassinavamajoniani (
:: <u>:</u> :			1985 83E88	ere erec		rž secou	p5 22
<u>r</u>		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	828 E88	600 640 640 640 640 640 640 640 640 640	Sports Sp	12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Consensus Section 172 Consensus Consensus
مقم خنين	2 1.c.d	0 000	S vicin	2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0000 0 0000 0 0000	g night	g nate

Figure 1A. Family A DNA polymerases—E. coli DNA pol. I (E.c. pol I)(1), Streptococcus pneumoniae DNA pol. I (S.p. pol I)(2), Thermus aquaticus DNA pol. I (Taq pol I)(3), T5 DNA pol. (4), T7 DNA pol. (6), and yeast mitochondrial DNA pol. (MIP1)(7).

	SCESSES.	4 <u>85558</u>	Control Margaria
002°	AVILLAM SERVICE OF THE SERVICE OF TH	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	400 William I Kirl I Habilit Warner Die Seif William I Kirl II Habilit Warner Die Seif William I Kirl I Habilit Warner Die Seif William I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif William I Kirl I Kirl I Habilit Warner Die Seif Will Will Warner Die Seif Will Warner Die
ř.	100 Marie 100 Ma	# 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	THE CONTROL OF THE CO
150	MERITEPPORTICALITY (ALL TITLE MERITEPPORTICALITY (ALL TITLE) (ALL	4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	The control of the co
<u>κ</u> -	Principality (A. 1994) (A.	CALLEGORICA CONTRACTOR OF THE	TASSESSION THE PROPERTY OF THE
8•	ANTERENT BESTELLEN TENENT TO SEE TO SECRETARY FRENCH. TO SECRETARY	HOTE IDECTIALATORE LI RODI VOGOSTATORISTOS DE LES SERENZAS DE LI DALLI TONICALI EL LA LOS SERENZAS DE LA LA LOS SERENZAS DE LA LOS SERENZAS DEL LA LOS S	
ĸ		ATEVELOFE TRAKELLING WOOD LUCK SEBARY GYTEVOEGY SUVAL GOOD LUCK SOUGHT	*
8,	######################################	A CONTRACTOR OF THE CONTRACTOR	TATE OF THE PROPERTY OF THE PR
х >		WPWAGGE SELECTIVE STATES OF THE SELECTIVE SELECTIVE STATES OF THE SELECTIVE	A CONTROLLED AND AND AND AND AND AND AND AND AND AN
			203/4-1-2511-5555-5-1121
œ	535 545 555 555 555 555 555 555 555 555	A Conservation	Address Addres







		33	ž ž		\$33 \$
0022				ikili Pineolovekaites Pin.	P (Wildiam With 1618
£.				Per Legition Constitution of Alexandron of A	
2150				A TRANSPORT OF THE STATE OF THE	OSEZ SEZ SEZ SEZ SEZ SEZ SEZ SEZ SEZ SEZ
X.*				ACTIVITY CONTROL OF THE CONTROL OF T	ZŢĢ ZZĢ ZZĢ ZZĢ ZZĢ ZZĢ ZZĢ ZZĢ ZZĢ ZZĢ
90 SZ				CHACLA SOET VIEW. Wheel Pemberwal worth Wheel Pemberwal worth Wheel Pemberwal worth Wheel Pemberwal Pemberwal Wheel Pemberwal Pemberwal Wheel Pemberwal	2200 (
K R				# EXVALXECTER ACO F. ED. # TYPESTYRE OCCUPATION F. F. D. L. # TYPESTYRE OCCUPATION F. F. D. L. # F.	ZZP
0502			Eccudintately and income		252 252 253 253 253 253 253 253 253 253
X.			WAGAILALEN WAGAINAL ITERICAN IN THE COMMINICAN THE	TOTAL THE LANGEST CONTROL TO THE PROPERTY OF T	TOTAL INTERIOR MANTER INTERIOR MANTER INTERIOR I
	-Strike	# 			
	A 64 64 64 64 64 64 64 64 64 64 64 64 64			1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	A COLOR OF STREET OF STREE

	767	
2609		0
क्ष्यं क्ष्यं क्ष्यं	<u>Limoptilikai way segsui sacii in saysai pichi ilusboy taakilakeni vilmaerabidairaah balarinadse piki misgyydayydogai tyati i 1sal i makeasolywoodiod</u>	2057 ZOPO ZOPO ZOPO ZOPO ZOPO ZOPO ZOPO ZOP
952		ZZŞO
88	(Crein V. Lamickas Lobitation	ZTZ
00\$2	WENTE CALLA LESSON TOWN VAN	ZÓD (KLIKASEKKI
24,3	FLEET WIN SECON SECON	2073 Odel Litterionia Trockelloid viet Fal
2650	PARTE SALES	2650
ສຸ	ĠŶĦĿĠĠŖŢĬŖĹĠĠŖŢĬŔĹĠŖŢĬŔĹĬĠĬĬŖĹĬĠĠĸĠŶŶĨĠĸĠŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶŶ	STOP STOP STOP STOP STOP STOP STOP STOP
	E.C. Design	Address Addres

Ven	1982
A CONTROL OF CONTROL O	2005 Spot

Figure 1B. Family B DNA polymerases—PRDI DNA pol. (9,10), \$429 DNA pol. (11), M2 DNA pol. (12), Adenovirus type-2 DNA-pol. (Adeno-2)(26), Adenovirus type-12 DNA pol. (Adeno-12)(28), Adenovirus plasmid DNA pol. (BCLK1)(32), Kluyveromyces lactis plasmid DNA pol. (BCLK1)(37), Kluyveromyces lactis plasmid DNA pol. (BCKL1)(30), Ascobolus immeraus plasmid DNA pol. (HSV-1)(19), Human cytomegalovirus DNA pol. (HCMV)(20), Epstein-Bart virus DNA pol. (GKL2)(31), hepes simplex type-1 DNA pol. (HSV-1)(19), Human cytomegalovirus DNA pol. (HCMV)(20), Epstein-Bart virus DNA pol. (FPV)(23), Vaccinia virus DNA pol. (VacV)(24), Autographa califomica nuclear polyhedrosis virus DNA pol. (FPV)(25), E. coli DNA varicella-zoster virus DNA pol. (YCX)(27), Yeast DNA pol. (FPV)(29), Yeast DNA pol. (16), Yeast DNA pol. (16), Yeast DNA pol. (16), Human DNA pol. (14), and Yeast Rev3 DNA pol. (18).

92 92 92 92	007	8 .	દેશ્યકે ક	200	-		7888 7888	<u> </u>	-	•	(35).
**			COLLYKY CACKED TANKER COLLYKY CACKED OF CACKED	800 PKYHERTTDENLEAFSFU DOVHRSEEFWELFADIP POOTHRSEEFWELFSDIP	MUDDENCE STATES		PLOOKKS PLOOKKS PLOOKKS	140U BIAVFKYHHALLYYAAYFTY GILLLKAHYPAEFHAANHA GILLKAHYPAEFHAANHA	1600 INALMIVEAREGE FLEXED	### 20 20	66 74
COLLASSMIGNEARGEGEDOGEDING I LASEGRANKE REGITNESGEDINAGYF I ENERCIC COL	375 XXXGHEALALTDHAVVQSFPOAN	K	WILCE THE VERBIE TOWNER TO TESTAVLIK DGLATIALVERAL GORALI TOFFILGE - LVK DGLATIALVERAL GORALI TOFFILGE - LVK	TTS WELLSGOOG AMPLEELPRYSPOOTHS EINAL MOGFLEDPERPRYSPOOTHS FIRST TO THE TOTAL TOT	975 975 GFBL POKT CPHCGTPLKKOGKO	r.	MODPTVIRM ODLY	1375 KKIKTHEPKAHAANVLHANEIAVFKYN OLVEKFAGTGEHKSHQANALVSTOTLALKA DLAEKFAGTGEKSHGANALVSTOTLALKA	1575 1575 1575 15110G - 1511PPFHS1PQLGT	COLDECTION CONTROL OF THE CONTROL OF	
3::	375 500 579 500 500 500 500 500 500 500 500 500 50		SSO LCRFLYPEFKHH. RLWTLCC. EPRFYHLRYHSDYSHIDGLAKT	750 EKI KEPWATGIYHYI.	PVAT-FORE	2. C	SELKTTHEDFHS INOMILIKLDILG TOFDKSOVEYAGLVKFDFLG TOFDKSOVEYAGLVKFDFLG	1350	1558	for Leepe of Eds 1 Acoupter	aloaralie faturyspsorl
	N25 Norwell was 1 per cape CE CEVE		525 Kasfdhgflavaykkllevekakappvíotli	725 FLEL VRDEKALKEL I ANITKLG	-[E[[R-R-E-17[MAVELA -[E[[R-R-E-E-1L 925	6-865-4-32-4-4-110-10-10-10-10-10-10-10-10-10-10-10-10	MOVIO TOFSPIOSPADATC APTKITOFACTOGEGKINA APTKITOFACTOGEGKINA - PCITOFACTOGEGKINA	1328 1328 1821 FVRECKGL. TPE 11 CGADHLRANGKKKPEENAKG	ILGGADHLRIAMKKINTERMIM 1525 NI ELAI ENCERGYSFOXVOL	SEGORIKAADONAKAERIGADA SEGORIKAADONAKAERIGATON 	Line Rose Epicoget Pork Tro
100 1000 110001754C1EELQQ15PP1	005	SLKKERAVIANIES I UND I VARIA TALIANI	SOO SZELIEDEREVIEDELIVANIASFONGFLIWAYTKILLEVEKAKIPVIOT	700 FDIASFYDYLEVOPEVRHILL	EECVAFTEAHPORY	FAVITIC SMELVANEL COSTING FPGFFL WEF COSKUNGVPV FPGFFL WEF COSKUNGVPV FFFL 1-5-0MGV	NOCE . TOWRITGONPOGIIV DWARKLEGYTRNAGKNAGOW DWABKLEGYTRNAGKNAGOW	1300 1300 VICCEDO INVIL I TOGLEPS L	11:EOM:1-0-L1-0-1-1	LEK INSGAPBRIGERIALIN LEK INSGAPBRIGERIALIN 1700	AREKYARGIAISI TORGIDOGL AREKYARGIAISI TORGIDOGL
190 1801.TTRL 105FAN LAVYTSSTEVODAEVSESTWODYSRCJEEL GGISPP I 181	Ŕ	CIKILDALISI I IDA VEREROVALISEE KERNYKARES I MOO I YROL YT AMAN	475 FAMPHIRPLEATITELTGTTOOMLODAPOVUDVIT	OUT CLASSICATION OF CLASSICATI	ILEGGRANDVCZCILRGMAKIVECVAFFA4PP0RY	, voeel pe ivear i ekëlksi i che ërikrapet , derle telov i moko erkrapet , derlo i elov i moko 	1075 KGVAGDMI IOGE - GAEIDBLV KTLAKAFEAEPOLPEIYEADEEVKALI KTLAKAFEAEPOLPEITEADEEVRALI	1275 1275 WAGARDELINNRICELSE	E	GGYFRELFOLCARTDÍRKLMRNY GGYFRELFOLCARTDÍRKLMRNY 1675	ŻYSFODSSCCLINI AREVTO IDE CAYSTODSSCCLINI AREVTO IDE
2:	0\$2	Wydovetrelkscatlei fki toyti	FAMPHRP	AST OS	CAYORGTOANGPIOSONALELIEUR CAYORGTOANGPIJERONAVELKEGL CAYORGTOANGPIERONAVELKEGL OSO	APLFPOEE	1 PPOPE	5	VAL FATOR OSGANON FOR HIGH	H VICE CALKOLOGES TEAT EARKE H VICE CALKOVEES TEAT DANS TO TO THE TO THE TEAT OF THE TEAT	Ex. Poolen S. F.* France Production (1195) E-M. Fidalokyodilemailiye E-N D-L T
SG 25 NGCL SWIMGOFGILLOGIWITOOF PHTF FENGE IKKLT I INKAKSHAFIF OF VELLE)	ສຸ	KESDEDAJA SGALVIGTO I KOMEE IRTLD SINDEERRI TVOGTVFDVETRELKSGRT	425 Vodovo i atriamiri Leeettuviõvettgi santot 1 ielaavavakalei 10kfu	650 SZ	TOTAGE TRANSCRILIDE IT, NITY LAMING TRANSITIES TO THE STANDER CLALADE FOR THE TOTAGE THE STANDER CLALADE FOR THE TOTAGE THE STANDER CLALADE FOR THE TOTAGE THE STANDER CLALADE FOR THE STANDER CLALADE FOR THE STANDER CLALADE FOR THE STANDER CLAUSE FOR TH	TRANCE INTERIORNEE WOOD IT PORT TOWN TOWN TOWN TOWN TOWN TOWN TOWN TOW	1050 201 2025 1050 1050 1050 1050 1050 1	1225 1225 1226	0571. 5271 STATE OF THE PROPERTY OF THE PROPER	MADO PO THE REAL TANDARD MELTINE WINDOWN THE REAL TOWNS OF THE REAL THE WAS CONTROL TOWNS OF THE REAL THE	PEGACY I ALE WALE WYTEREN IS ELLOS SELLEN FINALDY TOOL PEGACY I ALE WALE WATEREN IS ELLOS SELLEN FINALDY TOOL PEGACY I ALE WALE TO ALE WATEREN IS ELLOS SELLEN FINALDY TOOL PEGACY I ALE WATER SELLEN IS ELLOS SELLEN FINALDY TOOL PEGACY I ALE WATER SELLEN FINALDY TOOL PEGACY TOOL PEGACY I ALE WATER SELLEN FINALDY TOOL PEGACY
1 HEOLEVIEL	-					2000	971 FEGETOR		2000 2000 2000 2000 2000 2000 2000 200	25555 25555 25555 25555 25555 25555 25555 25555 25555 25555 255 255 255 255 2555 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255	100 100 100 100 100 100 100 100 100 100
•								A 20 1	-25 % -00 8		
<u>۔</u> آ	Consensus	S. Pol 16	Consent B.s. pol	Conservation	Consensus	6.8 pol 11. S.t. DNA E.	B.s. pol 115	S.T. O.	S. T. DRA E. S. T. DRA E. Consensus 1	B.s. pol 11 E.c. DuA E S.t. DuA E Consensua	B.s. pol 115 E.c. DMA E S.t. DMA E Conserrant

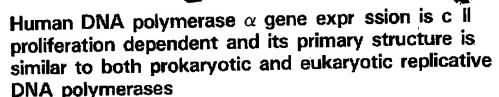
Figure IC. Family C DNA polymerases—Bacillus subtilis DNA pol. III (B.s. pol III)(36), E. coli DNA pol. III a-subunit (E.c. DNA E)(34), and Salmonella syphimurium DNA pol. III a-subunit (S.t. DNA E)(35).

22		i de la constante de la consta
200 FECHV. SOATHETH		
175 • • • • • • • • • • • • • • • • • • •	THE MAN TO THE PROPERTY OF THE	
מאמינים	NAME OF THE PROPERTY OF THE PR	
0\$1	ACCOUNT TO THE PROPERTY OF THE	
ğ,	THE CASE OF THE CA	
001	The first of the Language and the Control of the Co	
	TOTAL	
K,	22 22 23 24 24 24 24 24 24 24 24 24 24 24 24 24	
ã.		
ń *	MICHAELMAILLIAN MICHAELMAILUN MICHAELMAI	
·		
۵	A COURT OF THE PROPERTY OF THE	
	•	

Figure 1D. Family X DNA polymerases—Human DNA pol. β (38,39), Rat DNA pol. β (37), Human terminal transferase (Human TDT)(40), Bovine terminal transferase (Bovine TDT)(41), and Mouse terminal transferase (Mouse TDT)(41).

The EMBO Journal vol.7 no.1 pp.37 27, 1988

EXHIBIT C



Scott W.Wong, Alan F.Wahi, Pau-Miau Yuan1, Naoko Arai2, Barbara E.Pearson, Ken-ichi Arai², David Korn, Michael W.Hunkapiller¹ and Teresa S.-F. Wang

Laboratory of Experimental Oncology, Department of Pathology, Stanford Medical School, Stanford University, Stanford, CA 94305, ¹Applied Biosystems, Foster City, CA 94404 and ²DNAX Research Institute of Molecular and Cellular Biology, Palo Alto, CA 94304.

Communicated by A.Kornberg

We have isolated cDNA clones encoding the human DNA polymerase α catalytic polypeptide. Studies of the human DNA polymerase α steady-state mRNA levels in quiescent cells stimulated to proliferate, or normal cells compared to transformed cells, demonstrate that the polymerase α mRNA, like its enzymatic activity and denovo protein synthesis, positively correlates with cell proliferation and transformation. Analysis of the deduced 1462-amino-acid sequence reveals six regions of striking similarity to yeast DNA polymerase I and DNA polymerases of bacteriophages T_4 and ϕ 29, herpes family viruses, vaccinia virus and adenovirus. Three of these conserved regions appear to comprise the functional active site required for deoxynucleotide interaction. Two putative DNA interacting domains are also identified. Key words: primary structure/replicative DNA-polymerases/sequence similarity/structural gene/transcription

Introduction

Cell proliferation and the transmission and maintenance of error-free genetic information from one generation to the next are dependent on the mechanism of DNA replication. Genomic DNA replication is a complex and tightly regulated process involving the orderly coordination of many protein-protein and protein-DNA interactions (Kornberg, 1980, 1982). A key component of the chromosomal replication apparatus is DNA polymerase α , which is generally accepted as the principal polymerase involved in eukaryotic DNA replication (Kornberg, 1980, 1982; Campbell, 1986; Fry and Loeb, 1986). Many lines of evidence support this concept: its enzymatic activity positively correlates with DNA synthesis during cell proliferation (Fry and Loeb, 1986); all its specific inhibitors also inhibit DNA replication in vivo (Ikegami et al., 1978; Fry and Loeb, 1986); a mutant which is temperature sensitive for DNA synthesis has been identified as a DNA polymerase α mutant (Murakami et al., 1985); and monoclonal antibodies specific for DNA polymerase α inhibit DNA synthesis in permeabilized cells or when microinjected into nuclei (Miller, M.R. et al. 1985; Miller et al., 1986). The recent reports that DNA polymerase α plays a central role in SV40 DNA replication in vitro (Li and Kelly, 1984, 1985; Stillman and Gluzman, 1985; Wobbe et al., 1985; Murakami et al. 1986), a model system of eukaryotic DNA replication, further underscore the importance of this enzyme.

DNA polymerase α lacks the 3'-5' proof reading exonuclease which is required to remove mismatched nucleotides during DNA polymerization (Kornberg, 1980, 1982; Fry and Loeb, 1986). However, another eukaryotic DNA polymerase, designated δ , has been identified (Brynes et al. 1976), which possesses a 3'-5' exonuclease activity and. like polymerase α , is sensitive to the inhibitor aphidicolin (Lee et al., 1984). The relationship between polymerases α and δ is unknown. The identification of a cellular protein, proliferating cell nuclear antigen (PCNA), required for efficient viral DNA chain-elongation in vitro, and the discovery that this protein is able to stimulate DNA polymerase δ but not polymerase α activity, raises the question of whether there are two polymerases involved in eukaryotic DNA replication (Bravo et al., 1987; Perlich et al. 1987a,b). The recent finding of a cryptic proof reading 3'-5' exonuclease activity associated with the catalytic polypeptide of Drosophila embryo DNA polymerase α . when separated from other subunits (Cotterill et al., 1987). further stimulates interesting questions about the relationship between these two DNA polymerases.

Despite more than a quarter century of biochemical characterization of DNA polymerase α (Fry and Loeb, 1986) little is known about the regulation of the expression of this essential DNA replication enzyme, which nucleotide structural elements may be responsible for the cell-proliferation. dependent expression and whether its expression is the direct target of cascading biochemical events induced by growth factors or mitogens. In addition, nothing is known about the structure-function relationships of DNA polymerase a protein domains required for substrate recognition, or for the orderly coordination of protein-protein and protein-DNA interactions during chromosome replication. In an attempt to address these questions and to define the relationships between DNA polymerase α and δ , a near full-length cDNA of the human DNA polymerase α catalytic polypeptide has been isolated.

Comparison of the steady-state mRNA levels of quiescent cells stimulated to proliferate, or normal human cells compared to transformed cells, demonstrates that the previously reported increase of enzymatic activity during cell proliferation or transformation correlates with the level of steady-state mRNA. Analysis of the deduced primary structure of human DNA polymerase α with several viral DNA polymerases, yeast DNA polymerase I, Escherichia coli bacteriophage T4 and Bacillus phage \$629 DNA polymerase identifies six regions highly conserved among these polymerases. Three of these conserved domains appear to comprise th functionally active sites required for deoxynucleotide interaction and two other regions are postulated

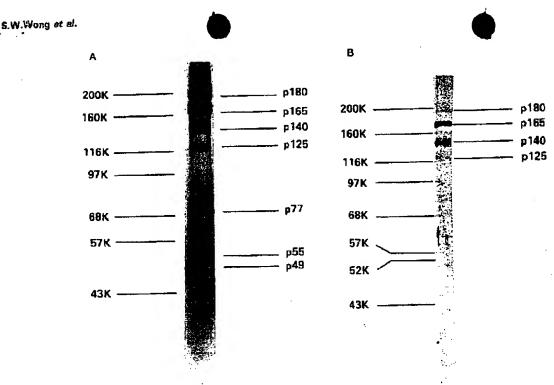


Fig. 1. SDS get analysis of human DNA polymerase α polypeptides. (A) Immunopurified human DNA polymerase α . (B) Isolated human DNA polymerase α catalytic polypeptides. 1% of the sample before and after HPLC separation as described under Materials and methods was analyzed on an SDS-8% polyacrylamide slab get and stained with silver.

Table I. I	Т9		T19		T23		T24		125		T264		T265	
Cycle no:		pmol		pmol		pmol	Res	pmol	Res	pmol	Res	pmol	Res	рто
1 2 3 4 5 6 7 8 9 10	S G Y S E V N L S K	20 54 58 11 41 46 24 30 8	A A Y A G G L V L D P	92 96 62 76 40 37 56 56 44 18 20	G P C W L E V K	24 26 5 8 25 14 19 5	N Y A F E I P D V P E	17 23 25 24 12 24 14 8 24 10	Y F D A E C A L E K	50 53 51 22 48 37 10 48 45 22	D T G N F V I G Q I L S	10.1 1.7 7.8 4.4 8.5 7.2 7.9 7.4 3.4 6.5 5.5 0.6	Q D N L T I D T Q Y Y L	3.9 0.7 2.5 5.3 0.7 4.3 3.8 0.7 2.7 1.6 2.2 4.2
12 13 14 15 16			K	19		٠	K	8			D Q S R	2.3 1.4 0.4 < <	A Q Q I	2.6 3.6 3.1 2.4

Shown are the PTH-amino acids observed at each sequence cycle. << indicates yield too low for quantitation.

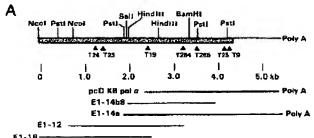
to be DNA binding domains, the presence of these conserved amino acid sequences among replicative DNA polymerases from phylogenetically distant species suggests they all may have evolved from a single primordial gene.

Results

Purification of the catalytic polypeptide and protein sequencing

The advancement of utilizing monoclonal antibodies specifically against human DNA polymerase α for immunoaffinity

purification has defined the protein structure and subunit components of this enzyme (Wang et al., 1984; Wong et al., 1986). DNA polymerase α contains (i) the catalytic polypeptide which, in vitro, is a family of large phosphopolypeptides of 180-125 kd, previously demonstrated by tryptic peptide mapping to be derivatives of the same primary structure (Wong et al., 1986); (ii) a 77 kd phosphoprotein of unknown function; and (iii) two polypeptides of 55 and 49 kd reported to be associated with DNA primase activity (Tseng and Ahlem, 1983) (Figure 1A). The catalytic polypeptides of human polymerase α were separated from associated



В GELEALATTEGESACIATEGIACETOTOGIACEGEACOACIETETETEAGATTEAGATTETOTTETATETETETECECESECEAGAAMMAATCAACAACGGGGGGGAGAACAAC 120 APAMBAETEMAN SECTIMACTEGETEMENTATION TO THE BASTEGAGGACTICACAGET OT TO TRANSPACT TEATCACACAGE TO TOTAL CACCACCT COTTEN SECTION OF THE CACCACT COTTEN SECTIO 115 155 TETECHTACTECTACECACTTECTTEABBAAMATTECTTCCCCTCTCTCTCAGAARCERCCTCCTTTAACTTCTGTTCTTTTAACTTCTTCAACTTCTTCTCCCCCTCTCTCAGAACTTCTCTCACCT CENCHETACHERACHOCHOGRAGICACCORCANTOCHOTTCANGATOCTECCTOCTATECCTECTATEC 275 76P 1#80 355 ACCARETOTOGTAFTTCTCTTTCCCHAAGTTTCCATTGAETCACCCCACCCCATCTTGAGCCCTTTCCTTCTTCTTCTTCATTGTTAAAATATETCACCCTTTACTTCCTTTCCCCCCCAAATCAA 132P CENAMOMOTATECTITICACATACCTRATECTICACAMAMICICACACACTICE ON NOT SEA NOT OLI FOR DELL'E TO TOGORICANCIANTETEGET FOLIE TENTETTENT ENERGEBARENTEN AMERICATETTENCT TENTETEN AND AND THE TOTAL CONTRACTOR TO THE TOTAL CONTRA 1929 ATATOCTTTTCAMENGTESTTERORMANGATGTERRESTTEREGTTECTECASCASAMARACACTGCTAERTTTTTTECTTERARAGTTRACAMITTCATCCTGTSATCATTGT TONTANTATT TO TEDERTY CAMET SEASON ACTACTICA CARLATINAT GTOTO CHARACTECT CATTOCT CAREAT A CONTRACT CAMETA CONTRACT CANCET CALCACT CARCATT CARC TCTANAGCTEANGESTTETANTCCCANTOCHARATATACAAAATATQTACASTGAATCTTCTCCAACCTETTATACCTETTTECAACCACCTCCAACCATCCAACATCCATTTTTCCACAT 2184 2124 CALACCACCTIATGCTEGAGGETTEGTTE TOCACCECAAAGITGC FITTIATGATAAGTTCATTTTGCTTCTGCACTTCAACACTCTATATCCTTCCATCATTCAACATTTA CEFERANCIEGTAFAAFERARANGAMETERAAGAGETAHTGAIACAGERAACATTAHATGIRGHEFITZITETTETETAYEACATTEERCHEFETTTERACETCAFACEGAACAG TATGTATEGTTEGGTEGGATTTTGGGTATARCAGATTTTACGCCAAACCAGTCCCTRCCTTGGTGACATACTACAAAGGAGGGATTTTEATGCATACAAACAGGTGGTACAAACCAGTCAA TETTEMETTATTETTATECHEATACHATTEATTATEATAAACACCAATACCACTATTETCAACTATTTATETTETCHATACCACTALLAGTECTA ACAMATTORICALATECTALATECTALATECTATTCCCA TERECTATTTERANTIALARECATTERANACEATTCCCLACCATTACCCTCATAMAMAGECTACCTCATCT ACATETTECCCCTCTGGATAAATTCTGAARAGGGAGAAACGTCAAAGCTBBAAGACTBBATATCTGTCTCTGTCTTCTGTCATCTGAATCCTCACTCCAAGTCAGACGGCCTATCCGCC CETTEMAGENTICAMTETICATCECCETACATETICACATATICACATATITATEATATATICETTICATCATTETICATETICATETICATETICATCATTECATICATCATTE AMANIFICATION THE TECHTURE TO THE TOTAL THE TO 4568 4888 4828 4828 4168 5468 5468

Fig. 2. Human DNA polymerase α cDNA. (A) Restriction map of human DNA polymerase α and overlapping cDNA clones. The stippled box represents the coding region of human DNA polymerase α and the solid line indicates the 5' and 3' non-coding region. Δ indicates the locations of each of the previously determined amino acid sequences. The five overlapping cDNA clones are pcD-KBpolα, EI-14b8, EI-14a, EI-12 and EI-19. (B) Nucleotide sequence and deduced amino acid sequence of human DNA polymerase α. Nucleotides are numbered at the upper right and amino acids at the lower right. Peptide sequences derived from immunopurified human DNA polymerase α preparation and used to design oligonucleotide probes as described under Materials and methods are underlined with dotted lines and labeled according to Table I. Amino acid number starts at methionine 1. * indicates the termination codon TAA.

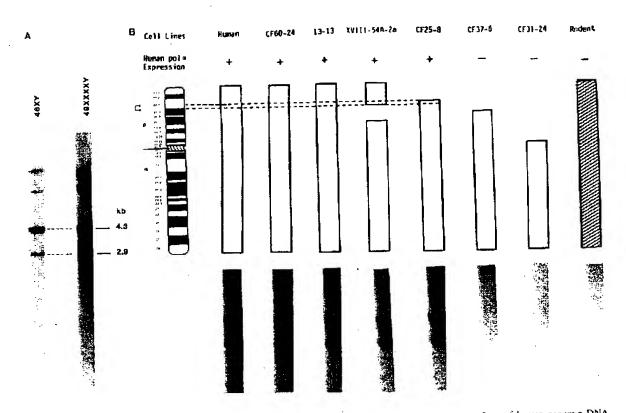


Fig. 3. Physical mapping of human DNA polymerase α cDNA on X chromosome. (A) Genomic Southern blot. 10 μg of human genomic DNA isolated from 46XY (1X normal male) and 49XXXXY (4X, from cell line GM1202A) were digested with EcoRI and hybridized as described under solated from 46XY (1X normal male) and 49XXXXY (4X, from cell line GM1202A) were digested with EcoRI and hybridized as described under the ferromasome is presented at the left. EcoRI digested genomic DNA, 10 μg each from human KB cells, rodent AKR thymoma cell line 3W5147 and somatic hybrids (Wang et al., 1985): CF60-24 containing human chromosome 4, 13, 17, 19, 21 and X; hybrid 13-13 containing only human X chromosome; hybrid clone XVIII-54A-2a containing the now interstitial deletion of most of the Xp21 band [46X del(X) (pter-21.3 :: p2i_1-qter)]; chromosome; hybrid clone XVIII-54A-2a containing the now interstitial deletion of most of the Xp21 band [46X del(X) (pter-21.3 :: p2i_1-qter)]; and the other three CF series hybrid clones containing X/autosome translocations: CF25-8 containing an X/13 translocation, retained der(X), Xqter-Xp21.1; CF37-6 containing an X/11 translocation, retained der(X), Xqter-Xp21.2; CF31-24 containing an X/20 translocation with dr(X), Xqter-Xp21.1; CF37-6 containing an X/11 translocation, retained der(X), Xqter-Xp21.2; CF31-24 containing an X/20 translocation with dr(X), Xqter-Xp21.1; CF37-6 containing an X/20 translocation with dr(X), Xqter-Xp21.1; CF37-6 containing an X/20 translocation with dr(X), Xqter-Xp21.1; CF37-6 containing an X/20 translocation with dr(X), Xqter-Xp21.2; CF31-24 containing an X/20 translocation with dr(X), Xqter-Xp21.2; CF31-24 containing an X/20 translocation and washing conditions were as described under Materials and methods.

subunits by gel permeation HPLC columns (Figure 1B). The separated catalytic polypeptides were pooled and treated as a single entity, digested with trypsin and fractionated by preparative reverse-phase HPLC. The amino acid sequences of seven peptides were determined (Table I) as described in Materials and methods. In all, the sequences of 85 amino acids were established and used to design single, long antisense oligonucleotide probes (Lathe, 1985) by which a near full-length cDNA clone of human DNA polymerase α was isolated.

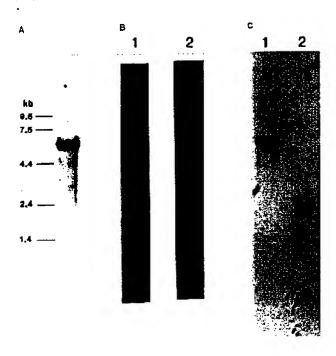
Primary structure of human DNA polymerase α

The 5433 nucleotides of the human polymerase α cDNA contain a single open reading frame coding for 1462 amino acids (Figure 2A and B). An in-frame initiator ATG codon flanked by nucleotides matching Kozak's criteria for a translation initiation site was identified (Figure 2B) (Kozak, 1981). All seven experimentally determined human DNA polymerase α peptide sequences listed in Table I are identified within this amino acid sequence (Table I, Figure 2A and B). Based on the deduced amino acid sequence, the estimated M_r of the recombinant polymerase α is 165 kd. Primer extension with two synth tic oligonucleotides corresponding to two separate regions of the 5'-end localizes the transcrip-

tion start site 295 nucleotides upstream from the putative translation initiation codon (data not shown).

Localization of the structural gene

The human DNA polymerase or gene was previously mapped by expression to a single genetic locus on the short arm of the X chromosome at the junctional region of Xp21.3 to Xp22.1 (Wang et al., 1985). The cDNA insert was shown to be X chromosome-linked by comparative genomic Southem hybridization with normal male DNA (46 XY) and DNA from a cell line of 4X (karyotyped 49 XXXXY) DNA (Figure 3A). Two EcoRI-digested genomic DNA bands are observed by hybridization with a PstI restriction fragment of pcD-KBpola (Figure 2A), both resulting in 1:4 ratio of signal intensity (Figure 3A). Using this PstI restriction fragment of pcD-KBpola cDNA clone, the chromosomal localization of the DNA polymerase α structural gene was analyzed directly by Southern hybridization of EcoRIdigested genomic DNA samples from a panel of human-rodent hybrids containing either an intact human X chromosome, different but overlapping regions of the human X chromosome and a hybrid clone with an interstitial deletion of the human X chromosome (Wang et al., 1985). Under conditions that exclude cross-hybridization to rodent



8197720061

Fig. 4. Steady-state mRNA analyses from quiescent cells stimulated to proliferate and comparison of normal and transformed cells. (A) Northern hybridization analysis of human DNA polymerase α mRNA. 10 µg of polyadenylated mRNA from early mid-log human KB cells was hybridized with 50 ng of 32P-labeled HindIII/BamHI 700-bp restriction fragment as described in Materials and methods. RNA sizes are given in kb and were determined by staining of parallel blot with RNA standard ladder marker from Bethesda Research Laboratory (BRL). (B) Steady-state mRNA from quiescent and proliferative cells. 25 µg of total RNA isolated from normal human lung fibroblast (IMR90), cultured in the 0.1% fetal calf serum (FCS) for 96 h to effect quiescence (lane 1) or stimulated to proliferate by activation with 10% FCS for 30 h (lane 2) were hybridized with 100 ng of 32 P-labeled anti-sense DNA polymerase α ribo-probe (5 \times 108 c.p.m./µg) as described in Materials and methods. Autoradiography was for 5 days-with intensifier screen at -70°C. (C) Steady-state mRNA comparison of transformed and normal human cells. 5 µg of poly(A)* mRNA from transformed cell line. MDA4, (lane 1) and from normal human proliferating tissue isolated from 20 week placenta (lane 2) were compared for relative abundance of steady state DNA polymerase a mRNA as described in Materials and methods.

Table II. DNA polymerase α activity in transformed and

Transformed cells	Units per 10 ⁷ cells	Non-transformed cells	Units per 10 ⁷ cells
293	19.0	TNHF	2.1
KB	14.0	GM1604	1.7
		IMR90 (midlog)	2.4
		IMR90 (quiescent)	UD

Units of DNA polymerase are defined as amol of labeled dAMP incorporated/h at 37°C in the presence and absence of 5 µg/ml aphidicolin. The results are averages of duplicate determinations. UD represents undetectable background value.

DNA, the structural gene for human DNA polymerase α was mapped precisely to the previously determined expression locus at Xp21.3 to Xp22.1 (Figure 3B).

Analysis of steady-state mRNA during cell proliferation and transformation

Characterization of this cDNA insert by Northern hybrid-

ization yields a single 5.8 kb band (Figure 4A) which is sufficient to encode a polypeptide of 165-180 kd. Northern hybridization of restriction fragments from each of the overlapping cDNA clones or the 5'-end of the near full-length cDNA clone all result in a single mRNA hybridization signal of 5.8 kb. These results indicate there is neither usage of multiple polyadenylation addition sites, generating mRNA of variable length from a single gene, nor various splicing events or processing occurring to generate multiple mRNAs.

The enzymatic activity and de novo protein synthesis of DNA polymerase a both correlate positively with cell proliferation (Bensch et al., 1982; Fry and Loeb, 1986; Thommes et al., 1986). To examine whether the transcriptional expression of polymerase a correlates with de novo protein synthesis and the expression of enzymatic activity, a parallel analysis of polymerase α steady-state mRNA levels from quiescent cells stimulated to proliferate was performed. Steady-state mRNA from normal human lung fibroblast cell culture (IMR-90), arrested in quiescent state (G0) by serum deprivation, and from cells activated to proliferate by serum stimulation were compared by Northern blot hybridization (Figure 4B). The steady-state message increases > 20-fold, 18 h after serum stimulation, 6 h prior to the peak of DNA synthesis. Meanwhile, the enzymatic activity per cell also increases ~ 10-fold (Table II). Therefore, the increase of DNA polymerase α steady-state message following the activation of quiescent cells to proliferate is similar to those observed with several other genes involved in DNA synthesis, that undergo transient expression.

A comparative study of message levels in normal growing human tissue and transformed cell lines was also performed. With equal amounts of poly(A)+ mRNA, no detectable polymerase a message was found in RNA isolated from a normal 20-week human placenta, in contrast to a readily detectable 5.8 kb signal from human breast carcinoma cell line (MDA4) (Figure 4C). Genomic Southern blots of normal human cells and transformed cells indicate equal gene dosage of human DNA polymerase α (data not shown), indicating that the abundance of the polymerase α message in transformed cells is not due to the amplification of the polymerase α gene.

Similarities with other DNA polymerases

Comparison of the primary amino acid sequence of human DNA polymerase α deduced from the nucleotide sequences of the cDNA clone to sequences derived from DNA polymerases of herpes, Epstein-Barr, cytomegalo, vaccinia, adeno-2 viruses (Gibb et al., 1985; Earl et al., 1986; Larder et al., 1987; Kouzardies et al., 1987), E.coli phage T4 (National Data Base Bank), Bacillus phage ϕ 29 (Yoshikawa and Ito, 1982) and yeast DNA polymerase I (Johnson et al., 1985) reveals several regions of marked similarity. Within a 472-amino-acid region of human DNA polymerase α (amino acids 609-1081) six regions are identified that contain extensive similarities among these DNA polymerases. The regions are designated according to the extent of similarity from I to VI with region I being the most similar (Figure 5A). In addition to these six highly conserved regions, the sequence spanning region VI of human DNA polymerase α , arnino acids 908-939, shares -41%similarity with T4 gene 46 protein, which is an exonuclease. The significance of these similarities is further underscored by the relative location of these regions within the respec. S.W.Wong et al.

. 3,7	trasdiid er m.						30	40
A	Human DNA pol a Herpes Simplex Cytomegalovirus Epstein-Barr Vaccinia Virus T4 DNA pol Adenovirus 2	609 437 379 350 212 185 384	VYEFP VYEFP	SEYE SELD	LLLGFM MLYAFF RIAKQL MLMEYI KKLPSI	L F F Q R Y A Q L I R D L S L E L T F D Y	PDIIVGHNIY PEFYTGYNII APAFYTGYNIN SVEIVTGYNVA YVYTFNGHN RPAIFTGWNIE LELYIVGHNI	GFELEVLLQR NFDWPFLLAK SFOLKVILTR NFOWPYILOR -FOLRYITHR GFDYPYIMNR
	Human ONA pol o Yeast ONA pol I Herpes Simplex Cytomegalovirus Epstein-Barr Vaccinia Virus T4 DNA pol Adenovirus 2	836 694 693 560 500 385 222 657	K A K Y V G - Y A V S Y R D G Y K F P Y	QGA T QGA T QGA T EGGK	V L O P T S V F E P E V V I Q P L S V F A P K Q	G - F H V I G - Y Y N G - F Y N K - M F S P - I A R	30 K F I L L L D F N S I N Y V L V M D F N S I N P V V V F D F A S I S P V L V V O F A S S R Y I M S F D L T S G E G M V F D V N S E P L Y V V D I C G	YPSIIMAHN YPSIIQAHN YPNVCIFGN YPSIIRQVN
	Human DNA pol a Yeast DNA pol I Herpes Simplex Cytomegalovirus Epstein-Barr Vaccinia Virus T4 DNA pol 429 DNA pol Adenovirus 2		1 C Y	STLL STMI ETLV	V P G G E Y F T P G E E H F G V V V S T F	V D P A D V R L A G L R P N R L E E E 1 P I H E Y I A	70 - E O G E - Q E Q - Q E Q - Q E I E V G G R - Y S V T L E N G V G E O Y E S F R L Y I N N Q - 12 - I T V G T A P K P S D E Y H D E O Y P L H I Q - Q A L D L Q G C - 27	I P E L P 0 - P I 0 E L P S V P R L F F V K - A T H R F V R - A G G V Y H F V K - K H C E P R I P H L I S C S P N G W H I R C E F - E P P P - F
	Tq exonuclease Human DNA pol o Yeast DNA pol l Herpes Simplex Cytomegalovirus Epstein-Barr Vaccinia Virus T4 DNA pol 629 DNA pol	ı	P S E H V R S V R H V H S E I	E M V O Q - E	GILPRE GVLPRL SULSIL SVLSEL SFLASE GIIPRL GIIPKE GIPTI	LANLVO LANLVO LNKWVS LTSWLA LRTFLA IAKVFF QI VPOERT	D L L E V G T L A E M R R K Q V K Q L M K C R R R E V K K V M K - M R K Q I R S R I P C Q R R A V R E C M R E K R K A I K K L L A A E R A R Y K K M L K K O R K O W K K K M F F K R S R F Y K G N E T V F P E W R C V A	10 K N K A L L T B Q O L M P J L L Q T E T D P A K R V Q C Q - D P V R E A V C Q - D P V R E M L A C E - D P Q R T A T S S T E K A L A E E M N A E S S G G S S R E Y V Q L N M M M S G S A E E Y V Q L N M M M M M M M M M
	Adenovirus 2 Human DNA polyeast BNA polyeast BNA polyerpes Simplex Cytomegaloviru Epstein-Barr Vaccinia Virus T4 DNA polye29 DNA polyepolyerus 2	I S	7 D 7 D 7 D	2 M Q Q L L L L L L L L L L L L L L L L L	130 ALKUTAN ALKUTAN ALKUTAN ALKUTAN ALKUTAN AN	SMYGCL SMYGCL SVYGFT AFYGFT AVYGFT SVYGLN	140 G:F S Y S R F Y A K G:Y V N S R F Y A K G:Y V N G L L P C L G:V Y N G M M P C L G:V A N G L F P C L G:F R N S A L Y S D G:N I H F R Y Y D L	150 P L A A L V T Y K G R P L A M L V T N K G R H V A A T V T T I G R F I A A S I T R I G R S I A E T V T L Q G R R N A T A J T L F G Q P Y L K E N G A L G F A A T L K G I T A G Q
	Human DNA pol Yeast DNA pol Herpes Simple: Cytomegaloviri Epstein-Barr Vaccinia Viru T4 DNA pol \$29 DNA pol Adenovirus 2	I x us	E I E M D M T M V G R - V N	LMAT LERT LERA ILYL JOWI LGEE IKSS	RQLAES' REYVH — ARFIK — KAFVE — ESVLN — ARKIN —	-28 — I [Y -52 — V [Y -27 — V [Y -43 — S V] Y -13 - — I A A -29 — I [Y	180 (G D T D S	R G L T A 899 R G L T P 923 R G F S E 768 D S Q O V 739 D K Y I E 631 T E I P O 468
	Human BNA pol Herpes Simple Cytomegalovir Epstein Barr Vaccinia Viru Adenovirus 2	rus eus	946 Y G 971 G A 815 T D	TKQE GKML SGLS GKTL PERI	LKGLDI IKGVDL MKGVEL MKGVEL NKGTSE AKGHAA	V R K T A C V R K T A C T R D V S E G L D Y O	1085 963 988 832 811 1107	

8197720061

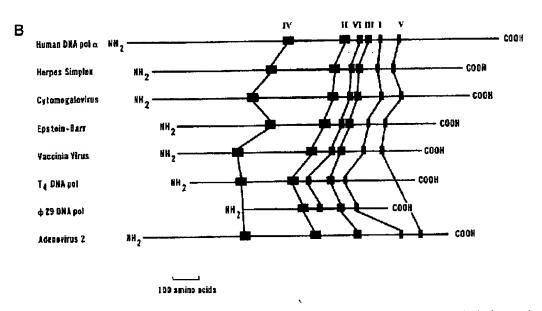


Fig. 5. Conserved regions of human DNA polymerases α and other DNA polymerases. (A) Amino acid sequence similarity between human DNA polymerase α and other DNA polymerases. Amino acid residues 609-1015 from human DNA polymerase α were aligned with amino acid residues derived from other DNA polymerases and T4 phage exonuclease gene 46 product; identical residues between human polymerase \alpha and other DNA polymerases in five or more sequences are boxed and similar amino acids between polymerase α and T₄ gene 46 protein (exonucl-ase) are boxed in dashed line. Gaps are indicated by dashes and extensive gaps are marked by the number of amino acids contained within the gap. The designated conserved regions are marked by dashed lines under the amino acid residues. Amino acids 998-1005 of human polymerase α are defined as region I; amino acids 839-878 are region II; amino acids 943-984 are region III; amino acids 609-650 are defined as region IV; region V and VI are amino acids 1075-1081 and 909-926, respectively. (B) Relative spatial arrangement of the conserved regions of DNA polymerases. Each DNA polymerase polypeptide is represented by a straight line with NH2 and COOH denoting the amino and carboxyl termini, respectively. The black bars represent the consensus sequences of each region. Similar regions of each polymerase polypeptide are aligned by vertical lines.

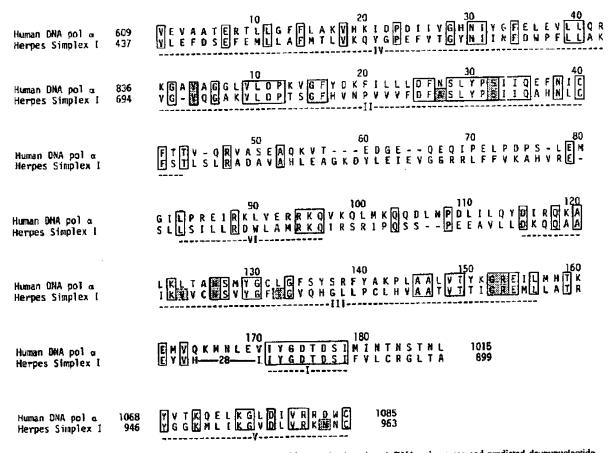


Fig. 6. Conserved sequences between human DNA polymerase α and herpes simplex virus 1 DNA polymerase and predicted de xynucleotide interacting domains. The six conserved regions between human DNA or and herpes simplex virus 1 were aligned. Identical amino acid residues are boxed. Amino acids that were identified in herpes simplex virus 1 mutants as single amino acid substitution are boxed in shade.

S.W, Wong et al.

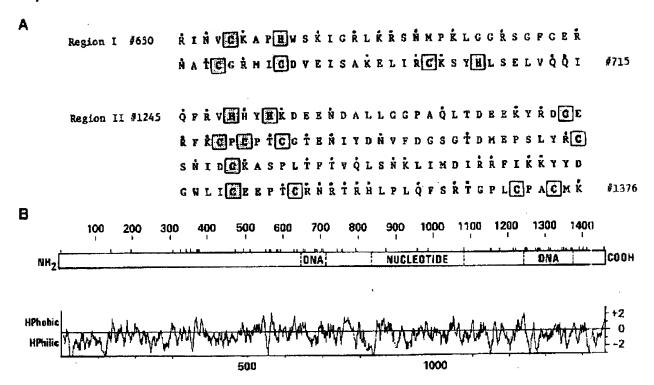


Fig. 7. Amino acid sequences of possible DNA binding regions and schematic summary of putative functional domains of human DNA polymerase α . (A) Possible DNA binding sequences. Two regions of cys/his-rich sequences of human DNA polymerase α are depicted, region 1 from amino acid 650-715 and region II from amino acid 1245-1376. Cysteine and histidine residues are boxed in shade and amino acids capable of interacting with the phosphate backbone of DNA in the possible loop regions are marked by \bullet . (B) Schematic representation of the putative functional domains of human DNA polymerase α and hydropathy plot.

tive polypeptides. The six regions are in the same linear spatial arrangement, IV-II-VI-III-I-V, on each polypeptide (Figure 5B). However, the distances between each consensus region in the polypeptides examined are variable.

Sequence comparisons of DNA polymerase β (Zmudzka et al., 1986; Matsukage et al., 1987), terminal transferase (Peterson et al., 1984, 1985) and E.coli DNA polymerase I (Joyce et al., 1982) reveal no significant similarity to the conserved regions described above, but a sequence similar to region II is identified in the dnaE gene product, the α subunit, of E.coli DNA polymerase III (Tomasiewicz and McHenry, 1987).

The presence of these highly conserved domains in replicative DNA polymerases from human to such phylogenetically distant species as bacteriophage T_4 and $\phi 29$ suggest that these DNA polymerases may all be derived from a common primordial gene.

Predicted functional domains

Conservation of these sequences is likely to reflect the need to maintain function. A detailed comparison was made of the amino acid sequences of human DNA polymerase α and herpes DNA polymerase (Figure 6). Extensive sequence similarity in all six regions are identified, with region I having the highest of 87.5%, followed by region II, 60%; region V, 57%; region III, 47%; region IV, 26%; and region VI, 10.5%. Mutations of herpes simplex virus conferring altered anti-viral drug sensitivity have been mapped to several of these conserved regions of the herpes DNA polymerase gene (Knopf et al., 1981; Coen et al., 1983; Quinn and McGeoch, 1985). Most mutants which demonstrate altered sensitivity

to the pyrophosphate analog phosphonoacetic acid also exhibit more resistance to the nucleoside analog aphidicolin (Coen et al., 1983; Larder et al., 1987). Sequence analysis of several of these mutants, derived from a single viral strain, confirms that all contain single amino acid substitutions within conserved regions II and III (Knopf, 1937; Larder et al., 1987; Tsurumi et al., 1987) and most recently, another mutant was identified having a single amino acid substitution in region V (J.S.Gibbs, H.C.Chiou and D.M. Coen, personal communication). Mutations conferring altered sensitivity to these drugs are inferred to be at the dNTP and polymerase mutants, regions II, III and V of human DNA polymerase α could be the essential catalytic domains required for dNTP interaction.

In eukaryotic cells, proteins involved in nucleic acid binding or gene regulation were found to contain cysteinehistidine rich sequences that are potential metal-binding domains which may play an essential role in nucleic acid binding and gene regulation (Miller, J. et al. 1985; Berg, 1986). Two regions containing such a motif are found in the DNA polymerase α sequence (Figure 7). One region, amino acids 650-715, contains the sequence Cys-X₃-His-X₂₇-Cys-X₄- $Cys-X_{11}-Cys-X_{3}-His$, where X represents amino acids other than cysteine and histidine, and Cys/His-Xn-Cys/His is capable of forming a tetrahedral box structure with an extended protein loop (Figure 7A). The extended loop between residues 659 and 685 contains many amino acids with side chains capable of interacting with the phosphate backbone of DNA (Ohlendrof and Mathews, 1983). Another cys/his-rich sequence within the carboxylterminus of human DNA polymerase α sequence, from amino acid 1245 – 1376, is identified (Figure 7B). This region contains the sequence His- X_2 -His- X_2 -Cys- X_4 -Cys- X_4 -Cys- X_2 -Cys- X_2 -Cys- X_4

Discussion

Genetic studies of yeast and several somatic cell lines implies that there are specific regulatory or restriction points in the growth cycle of the cell (Hartwell et al., 1974; Siminovitch and Thompson, 1978). Several genes involved in DNA synthesis such as thymidine kinase (tk) (Groudine and Casimir, 1984; Coppock and Pardee, 1987), thymidylate synthetase (ts) (Storm et al., 1984; Ayusawa et al., 1986) and dihydrofolate reductase (dhfr) (Farnham and Schimke, 1985) undergo transient expression in the cell cycle. DNA polymerase α is the principal enzyme that replicates chromosomal DNA. The steady-state level of polymerase α message increases when cells are activated to proliferate and correlates with the increase of enzymatic activity and de novo synthesis of antigenic protein. The concerted increase of these three parameters implies the regulation of the expression of this key DNA replication enzyme is at the transcriptional level. A gene that exhibits transient increase in expression during cell cycle or activation to proliferate could be a regulator of a restriction point, or the target of a cell-cycle or cellproliferation-specific regulatory signal. Thus far we have only analyzed the proliferation-associated steady-state message of polymerase a. The results imply that expression of polymerase a may be restricted to cells entering the cell cycle. The transient increase in polymerase a mRNA following the stimulation of quiescent cells could be the result of an activational event that renders the cells competent to induce transcription of DNA polymerase α and subsequently initiate DNA synthesis. To further understand the transcriptional regulation of this gene, investigation of the nuclear transcription and steady-state message of DNA polymerase α during activation of quiescent cells to proliferate, as well as within cell cycle, is necessary. The observation of significant amplification of steady-state polymerase α mRNA, denovo protein synthesis and enzymatic activity in transformed cells as compared to non-transformed cells, poses the question whether DNA polymerase α is a target for oncogene activation.

Recent studies from various eukaryotic systems indicate that, in vitro, the catalytic polypeptide of DNA polymerase α is a polypeptide of 180 kd (Campbell, 1986; Wong et al., 1986). The present amino acid sequence, deduced from the cDNA sequence, demonstrates that the catalytic polypeptide of human DNA polymerase α has a minimum mol. wt of 165 kd. The discrepancy between this value and the 180 kd polypeptide found in DNA polymerase α enzyme preparations suggests either there exists an additional 15 kd of coding sequence further upstream from the putative translation start site, or that a post-translational modification reduces its size. Compared to the yeast DNA polymerase I gene (Johnson et al., 1985) which encodes a biologically functional enzyme of 140 kd, a recombinant human polymerase α of 165 kd may represent the full-length polypeptide. Analysis of a recently isolated genomic clone containing the promoter region of human polymerase α will define the transcriptional and translational initiation sites as well as the primary translation product.

Structure – function studies of DNA polymerase α are prerequisite to understanding the mechanisms by which it interacts with other DNA replication proteins, dNTP and DNA substrates. Analysis of the primary amino acid sequence deduced from the cDNA has identified several structural features which are similar to other nucleic-acid-interacting proteins and DNA polymerases. The identification of cys/hisrich motifs defines putative DNA binding domains. The presence of highly conserved regions sharing similarity with other DNA polymerases and having similar, although not identical, spatial relationships suggests that these regions are needed to maintain essential function. Secondary stucture predictions for each of the three highly conserved regions, I, II and III, indicates that each composes a turn (cleft). Four regions, II, VI, III and I, are localized within a contiguous region that comprises ony 9% of the total length of the polypeptide. This close proximity prompts us to speculate that these domains may form a substrate binding site on the enzyme surface. Experimental evidence based on genetic marker transfer and rescue data of herpes DNA polymerase mutants (Knopf et al., 1981; Coen et al., 1983; Gibb et al., 1985; Quinn and McGeoch, 1985; Tsurumi et al., 1987) implicates regions II, III and possibly region V in substrate dNTP binding and ppi hydrolysis. Mutations clustered in locations such as region II and III suggest a role for these two sequences in substrate recognition or catalysis. Mutation in region V, which is > 100 amino acids apart from regions II and III, results in similar drug resistance phenotypes, suggesting polypeptide folding interactions that form substrate binding sites. The identification of a region II-like sequence in the α subunit of E. coli DNA polymerase III further substantiates the functional importance of this region. Thus far there are no regions I, IV and VI mutants isolated from herpes virus. These regions may be the critical domains required for interaction with other accessory proteins in DNA replication or for substrate interaction. The biological function of each of these conserved domains should be definable by site-specific mutagenesis and interchanging host and viral polymerase structural determinants.

It is interesting to note that all DNA polymerases containing these conserved regions identified in this study are replicative enzymes. In addition, except for DNA polymerase α and yeast polymerase I, all of the viral DNA polymerases and the two bacteriophage DNA polymerases have two enzymological activities; a DNA polymerizing activity and a 3'-5' proof reading exonuclease activity. Since errorfree DNA replication is an essential process for the survival of biological organisms, one might expect that the proof reading function would be conserved in this key chromosomal replication enzyme from prokaryotes to eukaryotes. This again raises the issue of the relationship between DNA polymerases α and δ . Does mammalian polymerase α , like Drosophila melanogaster polymerase a (Cotterill et al.) 1987), have an intrinsic but cryptic 3'-5' exonuclease activity in the catalytic polypeptide, detectable only when separated from other subunits? Enzymological characteriza tion of the functionally expressed polymerase a catalytic polypeptide should provide an answer to this issue.

The lack of these six conserved sequences in *E.coli* polymerase III, the chromosomal replicative DNA polymerase,

\$.W.Wong at al.



8197720061

suggests DNA polymerase α and E.coli polymerase III evolved from different ancestral genes. These six similar regions in replicative DNA polymerases are conserved to phylogenetically distant species; however, they are notably absent in several eukaryotic DNA polymerizing enzymes: DNA polymerase β (Zmudzka et al., 1986; Matsukage et al., 1987), terminal transferase (Peterson et al., 1985), retroviral reverse transcriptase (Kamer and Argos, 1984), or prokaryotic E. coli DNA polymerases I (Joyce et al., 1982) and III (Tomasiewicz and McHenry, 1987). This suggests that there is a class of DNA polymerases which are all replicative DNA polymerases, containing these conserved regions and sharing a primordial archetype.

Materials and methods

General methods

Preparation of plasmid, restriction enzyme digestions and agarose gel electrophoresis were performed as described by Maniatis et al. (1982). DNA probes were labeled with ¹²P by the methods of Feinberg and Vogelstein (1983), synthetic oligonucleotide probes were labeled by T₄ polynucleotide kinase as described by Maniatis et al. (1982). Anti-sense ribonucleotide probe was prepared as described (Melton et al., 1984).

Cell lines

KB is a human epidermoid carcinoma cell line; IMR90 is a normal human fetal lung fibroblast cell line; 293 is a human embryonic kidney cell line transformed with sheared adenovirus DNA; MDA4 is transformed human cell line propagated from human mammary carcinoma. These cell lines described above were from ATCC, Rockville, MD. GM1604 is a normal human fetal lung fibroblast cell line and GM1202A is human cell line containing 4X chromosomes, both these cell lines are from NIGMS Human Cell Repository, Carnden, NJ. TNHF is a normal human fibroblast primary culture from neonatal foreskin developed by Dr Eric Stanbridge of University California, Irvine, CA. All rodent-human somatic hybrids used in gene mapping were as previously described (Wang et al., 1985).

Isolation of human DNA polymerase α catalytic polypeptides DNA polymerase α antigen polypeptides from six 18 l cultures of human KB cells $(3.5 \times 10^5 \text{ cells/ml})$ were purified with a monoclonal IgG (SJK287)-Sepharose 4B column as described (Wong et al., 1986). The polypeptides were suspended in 0.36 M Tris-HCl, pH 8.6, 3.3 mM EDTA. 8 M urea and then reduced for 3 h at 37°C under N2 with 10 mM DTT. The reduced polypeptides were alkylated with 22 mM iodoacetic acid at 4°C for 1 h and dialyzed in 50 mM NH₄HCO₃, 0.01% SDS. The dialyzed, reduced and alkylated DNA polymerase a protein was lyophilized, resuspended in 100 mM NaPO4, pH 6.5, and 0.1% SDS and heated at 75°C for 10 min. These polypeptides were then purified by HPLC through two coupled gel permeation columns (TSK 3000, 7.5 × 300 mm) in 100 mM NaPO4, pH 6.5, and 0.1% SDS at a flow rate of 0.5 ml/min. The absorbance of the cluste was monitored at 280 nm. Fractions containing the 180-140 kd DNA polymerase a catalytic polypeptides were dialyzed in 50mM NH4HCO3 containing 0.01% SDS and lyophilized.

Poptide sequence analysis

Human DNA polymerase α catalytic polypeptides (500 pmol), isolated as described above, were resuspended in H2O and ethanol-precipitated twice to remove excess SDS from the samples. The polypeptides were then resuspended in 0.1 M $\mathrm{NH_4HCO_3},\,10$ mM $\mathrm{CaCl_2}$ and digested with 2 $\mu\mathrm{g}$ of TPCK treated trypsin at room temperature for 20 h. The trypsin digested peptides were first separated on an Aquapore RP300 (2.1 × 220 mm. Brownlee Lab) HPLC column equilibrated in 0.1% trifluoreacetic acid. A linear gradient from 0-60% acetonitrile was run over 45 min at 0.2 mi/min. Absorbance at 220 nm was monitored by Spectraflow 755 Variable Wavelength detector. Selected peptides peaks were further purified by an RP300 $(1 \times 100 \text{ mm})$ column equilibrated in 50 mM ammonium acetate, pH 6.5. A linear gradient of 0-75% acetomirile was run over 30 min at 0.08 ml/min and absorbance monitored at 215 nm. Each of the separated peptides was subjected to automated Edman degradation performed on a model 470A gas phase sequencer with on-line PTH amino acid anlaysis (Model 120A) (Hunkspiller et al., 1983).

Oligodeoxynucleotide probes

Single long anti-sense oligonucleotide probes were designed according to Lathe (1985) and were synthesized on an Applied Biosystems model 380A oligonucleotide synthesizer.

cDNA cloning

Ninety #g of poly(A)+ mRNA from early mid-log human KB cells was heated at 65°C for 1 min and loaded onto a 5.3 ml sucrose gradient of 5-25% containing 100 mM NaCl, 10 mM Tris-HCl, pH 7.4, 1 mM EDTA and 0.1% SDS. Centrifugation was carried out at 52 300 g for 25 h at 5°C and fractionated into 20 fractions. mRNA samples of each fraction oligonucleotide probes. Hybridization conditions used were $\alpha \times \text{SSPE}, 0.1\%$ SDS 100 µg/ml E.coli tRNA. Washing conditions were 2 × SSPE, 0.1% SDS. Temperature of hybridization and washing depended on the individual oligonucleotide probe used. Stringency of hybridization and washing of each individual oligonucleotide probe was based on $T_{\rm m}$ (melting temperature) and T_w (washing temperature) values estimated at >85% probe-target homology (Lathe, 1985).

Screening of 1 × 105 colonies of this size-selected library yielded a single distinct positive clone designated as pcD-KBpola, which hybridizes with oligodeoxynucleotides T264, T265 and T25 (Table I). Sequence analysis of pcD-KBpola (Figure 2) indicates that it contains a 2893-bp cDNA insert with an open reading frame of 1865 bp terminated by a stop codon and followed by a 1028-bp non-coding region. In this 1865-bp coding sequence there are four regions of deduced amino acid sequences that are perfectly homologous to the previously determined amin) acid sequences, T264, T265, T25 and T9 (Table D. The 3'-non-translated region contains several in-frame stop codons, and the consensus polyadenytation signal AATAAA (Proudfoot and Brownlee, 1976) 13 nucleotides upstream from the polyadenylation tail. This indicates that pcD-KBpola contains the 3' end of the cDNA for human DNA polymerase α . To extend this truncated cDNA clone the 5'-most restriction fragment of pcD-KBpola, PsiVHindIII, was used to screen 2×10^6 phage of a human pre-B cell cDNA library (E1 library) constructed in Agt10 (Cleary et al., 1986). The very 5'-terminal restriction fragments of the newly extended cDNA clones were used to further screen the Et library. The complete set of overlapping clones was sequenced in both directions as described (Dale et al., 1985) and reassembled.

Genomic DNA and RNA blot hybridization

Genomic DNA hybridization. Five or 10 µg of human genomic DNA were digested with EcoRI. DNA blot hybridization was carried out with 50 ng of ³²P-labeled Psti/Psti 700-bp fragment of pcD-KBpola (10° c.p.m./µg). Hybridization was at 6 × SSC, 50 mM NaPO₄, pH 7.0, 5 × Denhardt solution, 100 µg boiled and sonicated salmon sperm DNA. 50% formamide and 10% dextran sulfate at 42°C. The blot was washed in 0.2 × SSC, 0.1% SDS at 65°C.

RNA hybridization. Polyadenylated mRNA was analyzed on a 1% agarose get in 6.3% formaldehyde. Northern blot hybridization was carried out with 50 ng of ³²P-labeled restriction fragment of cDNA (10⁹ c.p.m./µg). Hybridization and wash were as described above.

Amino acid sequence of other ONA polymerases

Nucleotide sequence from yeast DNA polymerase I was determined from sequence analysis of a EcoRI/HindIII restriction fragment of yeast DNA polymerase I gene (Johnson et al., 1985). Other viral DNA polymerase sequences and T4 gene 46 sequence were derived either from published data or National Data Base Bank of the National Biomedical Research Foundation.

Acknowledgements

We thank Dr T.Yokota and Mrs K.Yokota for advice and construction of size fractionated pcD KB cDNA library; Dr M.L.Chary for the human pre-B cell cDNA library and cDNA cloning advice; A. Johnson for tissue culture assistance; Drs M.Synder and R.Davis for the yeast DNA polymerase I gene, Drs E.C. Friedberg and D.A. Clayton for helpful comments on the manuscript. This research was supported by a grant from the US National Cancer Institute (CA 14835) to T.S.F.W. and by a g.ft from the Donald E. and Delia B. Baxter Fund. S.W.W. is a predoctoral fellow and A.F.W. is a postdoctoral fellow supported by NIH training grants CA09302 and CA09151, respectively.

R ferences

Ayusawa, D., Shimizu, K., Koyama, H., Kaneda, S., Takeishi, K. and Seno, T. (1986) J. Mol. Biol., 190, 559-567.

Bensch, K.G., Tanaka, S., Hu, S.-Z., Wang, T.S.-F. and Korn, D. (1982) J. Biol. Chem., 257, 8391-8396.

Berg, J.M. (1986) Science, 232, 485-497.

Bravo, R., Frank, R., Blundell, P.A. and Macdonald-Bravo, H. (1987) Nature, 326, 515-517.

Byrnes, J.J., Downey, K.M., Black, V.L. and So, A.G. (1976) Biochemistry, 15, 2817-2823.

Campbell, J.L. (1986) Annu. Rev. Blochem., 55, 733-771.

Cleary, M.L., Smith, S.D. and Sklar J. (1986) Cell, 47, 19-28.

Coen, D.M., Fuman, P.A., Aschman, D. and Schaffer, P.A. (1983) Nucleic Acids Res., 11, 5287-5297.

Coppork, D. and Pardee, A.B. (1987) Mol. Cell. Biol., 7, 2925-2932. Conterill, S.M., Reyland, M.E., Loeb, L.A. and Lehman, I.R. (1987) Proc. Natl. Acad. Sci. USA, 84, 5635-5639.

Dale, R.M.K., McClure, B.A. and Houchins, J.P. (1985) Plasmid, 13, 31 - 40.

Earl, P.L., Jones, E.P. and Moss, B. (1986) Proc. Natl. Acad. Sci. USA. 83, 3659-3663.

Farnham, P.J. and Schimke, R.T. (1985) J. Biol. Chem., 260, 7675-7670. Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem., 132, 6-13. Fry, M. and Loeb, L.A. (1986) in Animal Cell DNA Polymerases. CRC Press,

Boca Raton, FL. pp. 13-60. Gibb, J.S., Chiou, H.C., Hall, J.D., Mount, D.W., Retondo, M.J., Weller, S.K. and Coen, D.M. (1985) Proc. Natl. Acad. Sci. USA, 82, 7969 - 7973

Groudine, M. and Casimir, C. (1984) Nucleic Acids Res., 12, 1427-1446. Hartwell, L.H., Culotti, J., Pringle, J. and Reid, B. (1974) Science, 183, 46-51.

Hunkapiller, M. W., Hewick, R.M., Dreyer, W.J. and Hood, L.E. (1983) Methods Engimal., 91, 399-413.

Ikegami, S., Toguchi, T., Ohasi, M., Oguro, M., Nagano, H. and Mano, Y.

(1978) Nature, 275, 458-459. Johnson, L.M., Snyder, M., Chang, L.M.S., Davis, R.W., and Campbell, J.L.

(1985) Cell. 43. 369-377. Joyce, C.M., Kelley, W.S. and Grindley, N.D.F. (1982) J. Biol. Chem., 257,

1958 - 1964. Kamer, G., and Argos, P.K. (1984) Nucleic Acids Res., 12, 7269-7283.

Knopf, K.W., Kaufman, E.R. and Crumpacker, C. (1981) J. Virol., 39, 746 - 757.

Knopf, K.W. (1987) J. Gen. Virol., 68, 1429-1433.

Kornberg.A. (1980) DNA Replication. W.H.Freeman, San Francisco, CA. Kornberg, A. (1982) DNA Replication, 1982 Supplement. W. H. Freeman, San Francisco, CA.

Kouzardies, T., Banker, A.T., Satchwell, S.C., Weston, K., Tomlinson, P. and Barrell, B.C. (1987) J. Virol., 61, 125-133.

Kozak, M. (1981) Nucleic Acids Res., 9, 5233-5252

Larder, B.A., Kemp, S.D. and Darby, G. (1987) EMBO J., 6, 169-175. Lathe, R. (1985) J. Mol. Biol., 183, 1-12.

Lee, M.Y.W.T., Tan, C., Downey, K.M. and So, A.G. (1984) Biochemistry, 23, 1906-1913.

Li, J.J. and Kelly, T.J. (1984) Proc. Natl. Acad. Sci. USA, 81, 6973-6977. Li,J.J. and Kelly,T. (1985) J. Mol. Cell. Biol., 5, 1238-1246.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

Matsukage, A., Nishikawa, K., Ooi, T., Seto, Y. and Yamaguchi, M. (1987) J. Biol. Chem., 262, 8960-8962.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zina, K. and Green.M.R. (1984) Nucleic Acids Res., 12, 7035-7056.

Miller, J., McLachlan, A.D. and Klug, A. (1985) EMBO J., 4, 1609-1614. Miller, M.R., Ulrich, R.G., Wang, T.S.-F. and Korn, D. (1985) J. Biol. Chem., 260, 134-138.

Miller, M., Seighman, C. and Ullrich, R.G. (1986) Biochemistry, 24, 7440-7445.

Murakami, Y., Yasuda, H., Miyazawa, H., Hanaoka, F. and Yamada, M. (1985) Proc. Natl. Acad. Sci. USA, 82, 1761-1765.

Murakami, Y., Wobbe, C.R., Weissbach, L., Dean, F.B. and Hurwitz, J. (1986) Proc. Natl. Acad. Sci. USA, 83, 2869-2873.

Ohlendorf, D. H. and Mathews, B. W.A. (1983) Rev. Biophys. Bioeng., 12, 259-284.

Okayama, H. and Berg, P. (1982) Mol. Cell. Biol., 2, 161-170.

Okayama, H. and Berg, P. (1983) Mol. Cell. Biol., 3, 280-289.

Perlich, G., Kostura, M., Marshak, D.R., Mathews, M.B. and Stillman, B. (1987a) Nature, 326, 471-475.

Perlich, G., Tan, C., Kostura, M., Mathews, M.B., So, A.G., Downey, K.M. and Stillman, B. (1987b) Nature, 326, 517-520.

Peterson, R.C., Chueng, L.C., Mattaliano, R.J., White, S.J., Chang, L.S.M. and Bollum, F.J. (1984) Proc. Natl. Acad. Sci. USA, 81, 4363-4367. Peterson, R.C., Cheung, L.C., Mattaliano, R.J., White, S.T., Chang, L.M.S.

and Bollum, F.J. (1985) J. Biol. Chem., 260, 10495-10502. Proudfoot, N.J. and Browniee, G.G. (1976) Nature, 263, 211-214

Quinn, J.P. and McGeoch, D.J. (1985) Nucleic Acids Res., 13, 8143-8163. Siminovitch, L. and Thompson, L.H. (1978) J. Cell Physiol., 95, 361-366. Stillman, B.W. and Gluzman, Y. (1985) Mol. Cell. Biol., 5, 2051 - 2060. Storm, R.K., Ord, R.W., Greenwood, M.T., Mirdamadi, B., Chu, F.K. and Belfort, M. (1984) Mol. Cell. Biol., 4, 2858-2864.

Thommes, P., Reiter, T. and Knippers, R. (1986) Biochemistry, 25, 1308 - 1314.

Tomasiewicz, H.G. and McHenry, C.S. (1987) J. Bacc., in press. Tscng, B.Y. and Ahlem, C.N. (1983) J. Biol. Chem., 258, 9845-9849. Tsurumi, T., Maeno, K. and Nishiyama, Y. (1987) J. Virol., 61, 388-394. Wang, T.S.-F., Hu, S.-Z. and Korn, D. (1984) J. Biol. Chem., 259. 1854-1865.

Wang, T.S.-F., Pearson, B.E., Soumalainen, H.A., Mohandas, T., Shapiro, L.J., Schroder, J. and Korn, D. (1985) Proc. Nail. Acad. Sci. USA, 82, 5270-5274.

Wobbe, C.R., Dean, F., Weisshach, L. and Hurwitz, J. (1985) Proc. Natl. Acad. Sci. USA, 82, 5710-5714.

Wong, S.W., Paborsky, L.R., Fisher, P.A., Wang, T.S.-F. and Korn, D. (1986) I. Biol. Chem., 261, 7958-7968.

Yoshikawa, H. and Ito, J. (1982) Gene, 17, 323-325.

Zmudzka, B.Z., SenGuputa, D., Matsukage, A., Cobianti, F., Kumar, P. and Wilson, S.H. (1986) Proc. Natl. Acad. Sci. USA, 33, 5106-5110.

Received on October 12, 1987